

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTA
TGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTTGG
TGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTC
ATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGT
GACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTT
TTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAAC
GTTATCATCAAATTAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGGACTTTCT
ATTGTGGCAAACCTCCAGAAAACAACCCTTTTGTCTGCACATGTAAGTGGAGCTGTGCTTACC
TTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAA
ATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCA
CTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAG
AAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAA
TGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATT
TCTTTACGGGTGGAAGCCAATTTACATGGATTAACCTCTATGACACTGCACCTTGCCCTATT
AACATGAACGAACACGGCTACTTTCCAGAGATATTTGATGAAAGGATAAAATATTTCTGTAA
TGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAAGTTGCTTATTCTTCTCTGAAATTT
TCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATAATCAGGAAACATGAAAG
AAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTATTAAAAACACCTATGCCT
ATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIA
AVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAV
LTFGMGSLYMFVQITILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDL
EQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPC
PINNERTRLLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGC
AGACCGTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGC
AGGAGCCTTCCTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCC
AGATACTATTTTTTGGATTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGA
TACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGCACCATGTTTG
AGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAA
TGAACCTGTGTGAATTCTGCTGATCCTGGTTTTCATGGTGCCTTTTTTACATTGGCTATTTTA
TTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTTATGGCTGA
CCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCCCAAAACATGGGA
TCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGAATCTCATGGCTCTTC
TTTCTGGATTTGGTGTCTGCAACTGCCCATACACTTACATGTCTTACTTCTCAGGAATGTGA
CTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGATCATAAGCA
AAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGGAAGTGCATAACAAAC
CATCAGGTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGGAAGTGAAAATCTTA
CTCTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAG
CTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTA
ATTTTCTTGTTACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATA
TTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATT
ATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATTTCTTCATTCTTGTGGAA
TAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTACCAAGTTCTTTTATGCCATCT
CTAGCAGTAAGTCCCTCCAATGTCATTGTCCTGCTATTAGCACAGATAATGGGCATGTACTTTG
TCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTTAGAATACCGCACCATAATCACTGAAG
TCCTTGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTGATGTGATCTTCCTGGTCAGCG
CTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAACAGGCACCAGAGAAGCAAATGGCAC
CTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTAGATATAAGAGG
GGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGCATTTTTT
CACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAGA
ACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGG
CTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGC
AAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAA
AGGTTATAGCTTTGCCTTGAGATTGACTCATTAATAATCAGAGACTGTAACAAAAA
AAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACT
TGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEIL
GVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWK
LGDPFPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILAL
ERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVD
ALEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGK
TDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNV
IVLLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFL
YLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTT
 CCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTG
 AGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTT
 CTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCC
 ATCCAGGTGTCATGCAGAATTATGGGGATCACCTTGTGAGCAAAAAGGCGAACCAGCAGCTG
 AATTTACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAA
 GTTGAAACAGCCTTGAAAGCTAGCTTTGAACTTGCAGCTATGGCTGGGTGGAGATGGATTC
 GTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCTGATT
 TGGAAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACT
 AACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCAAACCTGCAACA
 CAAACAACAGAATTTATTGTGAGTGACAGTACCTACTCGGTGGCATCCCCTTACTCTACAATA
 CCTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTG
 ATTTGTGTACAGAAGTTTTTATGGAAGCTAGCACCATGTCTACAGAACTGAACCATTTGTT
 GAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTA
 GTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGCTATGTCAAAGGTAT
 GTGAAGGCCTTCCCTTTTACAACAAGAATCAGCAGAAGGAAATGATCGAAACCAAAGTAGTA
 AAGGAGGAGAAGGCCAATGATAGCAACCTAATGAGGAATCAAAGAAAACCTGATAAAACCCA
 GAAGAGTCCAAGAGTCCAAGCAAACTACCGTGCATGCTGGAAGCTGAAGTTTAGATGAGA
 CAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCCTTACCCTGCCCCAGCTGGG
 GAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTTGGTTCCTAACTGGAATCAGC
 TCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGCCCTTCTCCTTATTGTAACCT
 GTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCTTCTAGCCTGGCTATGTCCTA
 ATAATATCCCCTGAGGAGAAAGGAGTTTTTGCAAAGTGCAAGGACCTAAACATCTCATCAGTA
 TCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTGAAAGCCAAGGAGTCACT
 GAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGACCCTTCTTCAGCTCTGAAAGAGAAA
 CACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTAAGAGCAAAAGAATGGCAGAAAAGTTT
 AGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAAGCTAAAATAAA
 GAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAACACAGACAGG
 GTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACTTACTT
 TTTCTGGTCTCTACCACTGCTGATATTTTCTCTAGGAAATATACTTTTACAAGTAACAAAAAT
 AAAAATCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATTACT
 CAGTAATTTGTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTCA
 AGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTA
 GAACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACT
 AATTTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCT
 TAATTTATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAA
 AAGTGCCATTAAACAAATGTATCACTAGCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGC
 AGAAATATTTGTGACAAAAAATTAAAGCATTTAGAAAACTT

FIGURE 6

MARCFSLVLLLTISIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRL
LGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQFAA
YCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPA
STSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAA
GLGFCYVKRYVKAFFFTNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTV
RCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCGAGCCCG
 GCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCCGAGCGCAACTCGGTCCAGTCGGG
 GCGGCGGCTGCGGGCGCAGAGCGGAGATGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGC
 TGGCGGCGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGC
 CCGGCCCCGGCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTG
 AGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAG
 AAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATG
 AGACCAACACAGACACGAAGGTTGGAATAAATACCATCCATGTGCACCCGAGAAATTCACAAGA
 TAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACG
 AAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCGAGCATGTACTGCC
 AGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGG
 ACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGG
 GCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCC
 AGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACC
 CCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGAT
 GCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTGGTGTATGTGTGCAAGC
 CGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCCCGAGAGAGGTCCCCGATG
 AGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGGAGGACCTGGAGAGGAGCC
 TGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCTGCACTGCTGGGAGGGGAAG
 AGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAAATAGCTAATTTATTTCCCCA
 GGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCCTACATCTTCTTCCCAGTAAGTTTCC
 CCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGCTCCCCCAGGCTGTTCTCCAGGC
 TTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAACTGCAGGAGCAGTTTGGCACCCCT
 GTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCGTTTGTCTACATGGCTT
 TGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTCTGATTGGTTTTGGGG
 AAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATGCAACAAATGAATT
 TTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGCAGATGAAATGT
 TCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCCTACCTCTGT
 GCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGGAGGGGGT
 CATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCCCAAGT
 CACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCTCTC
 CACTACCCACACCCAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGATT
 TTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAATAATTCTCACATCCCTCTAAAAGTAA
 ACTACTGTTAGGAACAGCAGTGTTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAAT
 GATATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTTGAAAGGTATATGACTGAGCG
 TAGCATAACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGA
 AATTTGCAAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACC
 GAGCAGGGCTGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACTCTACGCCACTC
 CACAAATGATGTTTTAGGTGTCATGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAA
 GTTTAAAGTTGCACATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACAT
 AAGTTGCATTTAGAAATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAA
 AAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQH
 KLRSVEEMEAEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQM
 VFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRMMLCTRDECCGDQL
 CVWGHCTKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLI
 TWELEPDGALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFME
 EVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
 327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCC
ACGGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCC
AAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTCTGG
ACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTAC
TGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACA
CTCCGTTACCACACTGGGTCATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAGATAGCC
CGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGC
ATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGC
AATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAATGCGTTC
ATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTG
TTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCCTGTCTTCTTTTTTTCTCCGGT
CGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCACTATTACTGGCTGCCCATC
ATGACCTCCATCCTGGGGGCTATGTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGT
GTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGG
CCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGAC
AACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGCACCCCACCCCCACCGT
CCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGTAAAAAAGGTTTTAGGC
CAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTGAGGCGGGCGGATCACCTG
AGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCCGTCTCTATTAAAAATACAA
AAATTAGCCGAGAGTGGTGGCATGCACCTGTATCCAGCTACTCGGGAGGCTGAGGCAGGAG
AATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAACC
TGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAAAGATTTTATTAAAGATATTT
TGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPHTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFW
TLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIA
RVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAF
MLLMRNIVRVVVLDDKVTDLFFFGLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPI
MTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYSKSLKILGKKNEAPPD
NKKRKK

Important features:**Transmembrane domains:**

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCC**ATG**GGGGCCTGCCTGGGAG
 CCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCT
 GCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCCTGG
 GGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCT
 GGGTGCTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCC
 TGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTT
 TCACCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAGAATGGGT
 TTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACG
 GCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCCTCTTCATCCTCATCC
 AGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGCAAGGCCGAGG
 AGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTTCTACTTGCTGT
 CGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCCAGCGGCTGCCACGAGGGCA
 AGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCCTGCCCA
 AGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCATCACCTCTACACCA
 TGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAA
 CCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATG
 CCCCAGCATTTGTGGGCCTCATCATCTTCCTCCTGTGCACCCTCTTCATCAGTCTGCGCTCCT
 CAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCA
 CACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACG
 GCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCTGGCCTCACTGCACGTCATGA
 TGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGATGATCAGCACGTGGACCGCCC
 TGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTCTACCTGTGGACCCTGGTAGCCC
 CACTCCTCCTGCGCAACCGCGACTTCAGC**TGA**GGCAGCCTCACAGCCTGCCATCTGGTGCCTC
 CTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTCCCCACACCAATCAGCC
 AGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTGCCCCCTGAGCCGGGCCTTCTAGTCGT
 AGTGCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCACACCCAC
 ACGGTGGAGCTGCCTCTTCTTCCCCCTCCTCCCTGTTGCCATACTCAGCATCTCGGATGAAA
 GGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACTCCCACC
 ACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCAGGGGACCCTGCC
 CCCTTCTGGACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCAGTGCG
 TGTAATAAAAAA

FIGURE 12

MGACLGACSLLSASCCLCGSAPCILCSCCPASRNSTVSRLI FTFFLFLGVLVSIIMLSPGVES
QLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPR
AAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQR
WLGKAEECDSRAWYAGLFFF TLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVS
IAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGY
ETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRA
FDNEQDGVITYSYFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICASWAGLLLY
LWTLVAPLLLRNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTTAGGGATGGT
 GAGGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAG
 AACGCTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCC
 ACACAACCTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTC
 AGGAGGACTTTCTGTTTGTGTTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATA
 GAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTAC
 TATCTTCATATTTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCA
 TATGCTGTGTGCAGACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCC
 TTTTACTAGCAAAAGTGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTG
 CCCATCATTTTCATTCATCCTTGCCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCT
 CAAGAAGCAGAAGAAGAAAACAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTT
 ATACCTGGTGGTCTTTCTGATGGTCAGTTTTATTCCTCCTGAATCCGAAGCAGGATCTGAA
 GAAGCTGAAGAAAAACAGGACAGTGAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTA
 AATGTGAAAAACCCTCACAGAAAGTCATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTG
 TCGACAGTAAAGTTGAAATGGTGACGTCCACTGCTGGCTTTATTGAACAGCTAATAAAGATTT
 ATTTATTGTAATACCTCACAAACGTTGTACCATATCCATGCACATTTAGTTGCCTGCCTGTGG
 CTGGTAAGGTAATGTCATGATTCATCCTCTCTTCAGTGAGACTGAGCCTGATGTGTTAACAAA
 TAGGTGAAGAAAGTCTTGCTGTATTCTTAATCAAAGACTTAATATATTGAAGTAACACTT
 TTTTAGTAAGCAAGATACCTTTTTATTTCATTCACAGAATGGAATTTTTTTGTTTCATGTCT
 CAGATTTATTTTGTATTTCTTTTTTAACACTCTACATTTCCCTTGTTTTTTAACTCATGCACA
 TGTGCTCTTTGTACAGTTTTTAAAAGTGTAATAAAATCTGACATGTCAATGTGGCTAGTTTTA
 TTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGACTTGCAAAAGGGGAAGAAAGG
 AATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTTTATCATGAAATCATGTTT
 TTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGCACAAAATGACTTAAACC
 ATTCATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAACTAAATTAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDL
LFVTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAI
ALTTAVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLPQEAEENRLLIV
QDASERAALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCCGACCCGCCAGGAAAGACTGA
GGCCGCGGCCTGCCCCGCCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAG**ATGT**GCT
CCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCT
GCCCATCCGGCTGCCAGTGACGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCA
CGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCA
TGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACC
AGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGA
CGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGGCCTCGAGCGCC
TCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCC
TCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCC
TGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTG
CCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCA
GCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTG
TGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACACCCGCATTGCCAGC
TGCGGGCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCC
TGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCC
GCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCTGGGTGCGCGAGAGCCACG
TCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCAAGAACGCTGGCCGGCTGC
TCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAGCCACCACCACACAGCCACAGTGCCCA
CCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCTTCTAGCTTGGCTCCTACCTGGCTTA
GCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCTCCACTGCCCCACCGACTGTAGGGC
CTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCAATGGGGGCACATGCCACCTGG
GGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTTCACGGGCCTGTACTGTGAGAGCC
AGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCACGCCGAGGCCACCACGGTCCCTGA
CCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGCAGCGCTACCTCCAGG
GGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCGGGCCCTGATAAGC
GGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACACGGTCACCCAGCTGCGGCCCA
ACGCCACTTACTCCGTCTGTGTATGCCCTTGGGGCCCCGGGCGGGTGCCGGAGGGCGAGGAGG
CCTGCGGGGAGGGCCATACACCCCAGCCGTCCACTCCAACCACGCCCCAGTCACCCAGGCCC
GCGAGGGCAACCTGCCGCTCCTCATTGCGCCCCCCTGGCCGCGGTGCTCCTGGCCGCGCTGG
CTGCGGTGGGGGCAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGGACA
AAGGGCAGGTGGGGCCAGGGGCTGGGCCCTGGAACCTGGAGGGAGTGAAGGTCCCCTTGGAGC
CAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGC
CACTCATGGGCTTCCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACAT**CTAAG**
CCAGAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTG
CTGCCACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGAC
CACAGCTGGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGC
TGACGAGCCCTAACGTCCCCAGAACCAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGC
AACGTGCAGTCCCTGGGCACGGCGGGGCCCTGCCATGTGTGGTAACGCATGCCTGGGTCTCTGC
TGGGCTCTCCCACTCCAGGCGGACCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAG
GGAGAGCGGGTAGGCGGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAA
ACTGGAAGGAAGATGCTTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAG
ATCCTTTCCCATTTATTCTGGGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTG
TAAGACAAACGATGATATGAAGGCCTTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G
 I T M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L
 E R L Y L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P G I L
 D T A N V E A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I
 A Q L R P E D L A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E
 S H V T L A S P E E T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T
 W L S P T A P A T E A P S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C
 E S Q M G Q G T R P S P T P V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P
 D K R L V T L R L P A S L A E Y T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A V H S N H A P V T
 Q A R E G N L P L L I A P A L A A V L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P
 L E P G P K A T E G G G E A L P S G S E C E V P L M G F P G P G L Q S P L H A K P Y I

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

[illegible]

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSKTTLTSDSVKDHTTAGRVVAGQIFLDS
EELSESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEP
CHFPFLLFDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAARRQMGEAEMMYQT
GMKILNGSNKKSQKREAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEEKLTEE
GSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMLVLSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATTCTGCAGTGGAAATTCATGAACTAGCAAGAGGACACCATCTTC
TTGTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGC
TAGGCCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAACTAAAGATTGAAGACA
TCACTGATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAA
CTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTT
TAAAGGCAGAAACCTCAGAGAGACTTCGTAAGTGTGCTTCTGGATGTGACCGACCCAGAGAATG
TCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCA
ATAATGCTGGTGTTCCTGGCGTGTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAG
AACCTATTGAAGTGAACCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCA
AGAAAGCTCAAGGGAGAGTTATTAATGTCTCCAGTGTGGAGGTCGCTTGCAATCGTTGGAG
GGGGCTATACTCCATCCAAATATGCAGTGGAAAGTTTCAATGACAGCTTAAGACGGGACATGA
AAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAACAAACTTGGCAGATC
CAGTAAAGGTAATTGAAAAAACTCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAAC
AATATGGAGAAGGTTACATTGAAAAAGTCTAGACAACTGAAAGGCAATAAATCCTATGTGA
ACATGGACCTCTCTCCGGTGGTAGAGTGCATGGACCACGCTCTAACAAGTCTCTTCCCTAAGA
CTCATTATGCCGCTGGAAAAGATGCCAAAATTTCTGGATACCTCTGTCTCACATGCCAGCAG
CTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGGCTAATCCCAAGGCAGTGTGAC
TCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGATTTCAAGAACACATCTCCT
TTTCAACCCCATTCCTTATCTGCTCCAACCTGGACTCATTTAGATCGTGCTTATTTGGATTGC
AAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTGCTCAAGTTTTCTTTGAAAAG
GAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCTGTATTTAGGCTTTGCCTGCTTGG
TGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGATCTTTACCGTGGCCTGCCCCA
TGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGTTAAGTATCATCTTATCTAAATA
TTAAAAGATAAGTCAACCCAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTE
SGSTALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLT
LEDYREPIEVNLFGLISVTNLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDS
LRRDMKAFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKG
NKSIVNMDLSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELAN
PKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGG
CGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGT
AAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATAC
AATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGA
AGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTA
CAAATTCCGTCGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTT
GCAGGAGCATTTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGA
AAGCTGCTCTACTCATCGACTGGAACATTCTTATATAAACCTCAAAAAGGACTTTTTTCACAG
GGTACCTTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTATATAAACTGTATCAGG
TTCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGA
AGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATT
AAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGT
AAACAGATTAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAA
GAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTGAGGCATTACGGACCTTTTTTCC
AAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAG
CTGTA ACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAATGGTAGAACACACTGA
CATTCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAAGCATAAAGCCTTAGACTTAGA
TGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATAACAAGACAAACGATCTAAAGCAAA
TACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCCCAGAAACAGATGAAGAAAT
TGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACATTTTGATCCTTTTAACCTTA
CAAGGAGATTTTTTTTATTTGGCTGATGGGTAAAGCCAAACATTTCTATTGTTTTTACTATGTT
GAGCTACTTGCAAGTAAGTTCATTTGTTTTTACTATGTTTCACCTGTTTGCAGTAATACACAGAT
AACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCAGATGCTTTTATTTCCAAAC
CTTTTTTTCACCTTTCACCTAAGTTGTTGAGGGGAAGGCTTACACAGACACATTCTTTAGAATT
GGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTTAGGGAAGACAAGTC
AGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATTGAGACCATGTCTA
TTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAAGAAATTTATAT
GAAAATTTATCTGAGTCATTAAAATCTCCTTAAGTGATACTTTTTTAGAAGTACATTATGGC
TAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCTAAAATTT
AAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLGGEVKGEAKNSITDSQMDDVEVVYTIDIQK
YIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFERRHSDQIMTFRERLLHKNLQEHFS
NQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMST
GFSRAVQTHSSKFFEEGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKR
EIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNH
HLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLDQTQDKRŠKANTGSSN
QDKASKMSSPETDEEIEKMKGFGEYSRSPTF

Important features:**Signal peptide:**

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCC
CAAGCAGCGCGCAGCGAACGCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCTCC
TTCCCCGCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCCGCGAAACCCCGAGGTCACCAGCC
CGCGCCTCTGCTTCCCTGGGCCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCCGCGCCTGGCACCGGGG
ACCGTTGCCTGACGCGAGGCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCACCACT
CCAACCTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAGCGCCGC
TTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGGCCCGCACCATGGCACGGTTCGGCTTGCCCGCGCTT
CTCTGCACCCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGA
CGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATC
TGTCCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAA
AGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTGATGAATTCTTC
AAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAA
AATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAA
ATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGTACCCTTTACAGAT
GAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCTCGCAAATTGAAGCTC
CAGGTTACTCGTGCTTTTGTAGCAGCCGTACTTTGCTCAAGGCTTAGCGGTTGCGGGAGATGTGCTGAGCAAG
GTCTCCGTGGTAAACCCACAGCCAGTGATCCCATGCCCTGTTGAAGATGATCTACTGCTCCCACTGCCGGGGT
CTCGTGAAGTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGCTGTTTGGCCAACCAAGGGGATCTC
GATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGAGGGTCCTTTCAACATTGAA
TCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGATAATAGTGTTCAAGTGTCT
CAGAAGGTTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGTTCATCTCTGAAAGT
GCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACTAGTTTGGACCGA
CTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCCTCCCTTCGAGCAACGTTTGAAC
GATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTACCTGTTT
GCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACCAGACATA
CTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCA
GAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTC
GGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCTGTTATGCAGAGAGAGTGGAGATTAATTCTCA
AACTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAAGGCACCAGTTATCACTTTTCTACCATCCTAGTGACTTTGC
TTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAAGTGCTGACTTTGTTTC
TCATTAGTTTTGGGAGGAAAAGGGACTGTGCATTGAGTTGGTTCCTGCTCCCCCAAACCATGTTAAACGTGGCT
AACAGTGTAGGTACAGAACTATAGTTAGTTGTGATTTGTGATTTTATCACTCTATTATTTGTTTGTATGTTTTT
TTCTCATTTTCGTTTGTGGGTTTTTTTTTCCAACGTGATCTCGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTT
CTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTTATCATGTTATC
TTATTAAAGAAAAAGCCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQG
 STCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVK
 TYGHLYMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLEC
 VSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMI
 YCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPI
 DVKISDAIMNMQDNSVQVSQKVFQGCQPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAA
 GTSLDRLVTDVKEKLKQAKKFWSSLPNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGL
 ANQGNNEPVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGEGSGSGCEY
 QQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:**Signal peptide:**

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTG
ACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGC
AACTTACAGCTGCACCGACAGTTGCG**ATG**AAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGC
CACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGG
ACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATT
GGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCC
CCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAA
ACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGC
CTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAG
CACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCAT
TCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTCTCTAGTGTCTT
CTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAA
GATTCCAGGAACTGTAGCTTCCTAGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTA
GCAAACAGAAGTCAATAAATATTTTTTAAATGTCAAAAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRK
FMTVSGLPKKQCPCDHFEGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTC
AGAGCTGGTCTGCCATGGACATCCTGGTCCCACTCCTGCAGCTGCTGGTGTGCTTCTTACCC
TGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCT
ACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCT
TCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCG
GAACCGGAGCCAACTTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATC
CCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGT
TTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCT
GCACTCTGGTGTGCTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTAC
TGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCT
TCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTGCTGCCTACCA
GAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGGAACGACAGCCCC
CTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCC
CAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGC
CTATCTATCTTCCACTGAGAGGGACCTTAGCAGAATGAGAGAAGACATTCATGTACCACCTACT
AGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATCCCGCCTTCGACAGTGA
AAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCCTGTTGTATCCTCAACTG
CAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTCCCTTTCTTTCGTTCCCAT
GGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGTCTCTAGGAACTGGTCACAA
AAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCTCTCCCCACTACCACCTTCTT
CCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGATGCCAGAGCAAGACTCAAAGAG
GCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAAACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIK
GLTGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAP
GEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQ
VFPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFSSKA
LICSFPSLQLEQATHQPIYLPPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCT**ATG**CTGCTGCTAACGCTGCTGCTG
CTGCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCT
TCTGCCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTG
CCTCATCGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGG
GATGGC**TAA**GAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGC
CTTATTTAAAGTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTT
CATTAAGAAATTGTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATG
TCTTTATAATAAACAGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180

FIGURE 30

MLLLTLLLLLLLLKGSCLWGLVGAQKVSSATDAPIRDWAFPPSFLCLLPHRPAMTCSQAQP
RGEGEKVG DG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAG
TTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACT
CCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATCA
TGTCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGA
TGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCAATTGGTTATTTTGG
GATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCA
TAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTATCCACAG
GCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTG
AGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGT
GGACATTTGCCATCCTCATTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAA
CTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTTCGGGCATTC
GGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCC
AGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTC
CTGATCATCCCATCCTTTCTGCTCTCTCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGA
AAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCATGTACATGCAAACG
CACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGATGCTGCTACTGCTGTT
TCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATATACTACAACTGCTATTA
ATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAATCTTGTCCTCAAGAACTCAAGTC
ACTTTACATCTATTAAGTCTTTGGAGACTTCATAATTTTTCTAGGAAAGGTGTTAGTGGTGT
GTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCGGGCATTCCAGGTGTGGGCAG
TCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCCATAGTTTTTTATCTGTGTTTG
AAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTGATCTGGAAACAAATGATGGATCGT
CAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTAAAAAGGAGCAACAAATTAA
ACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGAGGGAACAGAACTCCAGG
CCATTGTGAGAT**TAG**ATACCCATTTAGGTATCTGTACCTGGAAAACATTTCTTCTAAGAGCCA
TTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTTTTTAAAGACCTAA
TAAACCCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTDNLS
IELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPL
WTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILAC
QQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQN
ALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSS
HFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVF
ETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQ
AIVR

Important features:**Signal peptide:**

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

[illegible]

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCP
 AGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSL
 PRWRESFIVLESKPKKGVTPSALTYSSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLA
 VTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDP
 SGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQL
 LADVAQALDIGPAGPLMGVVQYGDNPATHTNLKTHNSRDLKTAIEKITQRGGLSNVGRAISF
 VTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVV
 EPNFANKAVCRTNGFYSLHVQSWFGLHKTLQPLVKRVCDTDRLACSKTCLNSADIGFVIDGSS
 SVGTGNFRTVLQFVTNLTKEFEISDTRIGAVQYTYEQRLFGFDKYSSKPDILNAIKRVGY
 WSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKGVITYAIGVAW
 AAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
 431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAA
GAAATTGCCAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTT
TAATTAAGCATGGAATACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCCA
CAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCAC
GTGGTGCTCTCCGACTACTACCCCGAGTGTAAGAACCCTTCGGCTCGCGTGCTTCTGAGCTG
CTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCCTGAGATCCCTCAAA
TGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCTTCCCCAC
TACAATGTGATAGAACGCGTGAACCTGGATGTACTTCTATGAGTATGAGCCGATTTACAGACAA
GACTTTCACTTCACACTTCGAGAGCATTCAAAGCTCTCATCAAAATCCATTTCTGGTCATT
CTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAA
AAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAG
GAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGA
CAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGCATTGAGGTGGGTA
ACTGAGTTTTGCCCCAATGCCAAGTACGTAATGAAGACAGACACTGATGTTTTTCATCAATACT
GGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTTTTCACAGGTTATCCT
CTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATATTTCTTACCAGGAGTAT
CCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATAATGTCCAGAGATTTGGTG
CCAAGGATCTATGAAATGATGGGTACGTAACCCATCAAGTTTGAAGATGTTTATGTCGGG
ATCTGTTTGAATTTATTAAAAGTGAACATTCATATCCAGAAGACACAAATCTTTTCTTTCTA
TATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCC
AAGGAGATCATCACTTTTTGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAACCTTAC
ATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTG
TGGAATTTTATGAGGAGGTCAGTGTGCTGGCTTACACTGAACTGAACTCATGAAAACCCA
GACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGGCCCTTCAAAGATGATATGTGG
AGGAATTAAATATAAAGGAATTGGAGGTTTTTGTCTAAAGAAATTAATAGGACCAACAATTTG
GACATGTCAATTCTGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCACTAGG
CTGTAAAAACAAAACAATGTAGAGTTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGT
GTATATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCA
CTGAAGTTATACTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTT
GCAGTATTTACAGTTATTATTATTTAAATTAATTTCACTTTGTGTTTTTAAATGTTTTGAC
GATTTCAATACAAGATAAAAAGGATAGTGAATCATTCTTTACATGCAACATTTTCCAGTTAC
TTAACTGATCAGTTTATTATTGATACATCACTCCATTAATGTAAAGTCATAGGTCATTATTGC
ATATCAGTAATCTCTTGGACTTTGTTAAATATTTTACTGTGGTAATATAGAGAAGAATTAAAG
CAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDF
HFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEEKD
KMLALSLEDEHLLYGDIIHQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGN
LVKYLNLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPR
IYEMMGHVKPIKFEDVYVGICLNLLKVNIIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKE
IITFWQVMLRNTTCHY

Important features:**Type II transmembrane domain:**

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

[illegible]

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVG
 YTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCM
 RCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDN
 RDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFGHAIYEEITACSSSPCFHDGTCVL
 DKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFC
 NNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLHQLYSAA
 FSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRRTGKWSGRAPS
 CIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAH
 CVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLL
 DKARISTRVQPICLAASRDLSTS FQESHITVAGWNVLADVRS PGFKNDTLRSGVSVVDSLLC
 EEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMGLVSWSY
 DKTCSHRLSTAFTKVL PFKDWIERNMK

Important features of the protein:**Signal peptide:**

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
 474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTA
ATCTGAAGGTTCTCAGTCAAATTTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTT
GGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGC
TTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATT
CAGAACTCTGTAAAGGTGCCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAG
ACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAG
ACGAGCCTGGCCTAGACAACCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACT
CTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAA
ATCGAGCTTTGAGTGTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAA
ATTCTGAAAACACCACTGCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCA
TCAAGATCAATCGAGTAGATCCCACTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCC
ATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTC
TAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGG
TGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCGCGCAGCAGGAACAATGGACAGGCCCCGGATGCCTACAGAC
CCCAGATGACAGCTTTTATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAATAAACTGGTGCGCA
AGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGACATGGTCAGCTTGAGG
AGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCCAGAAAGTGGCGCTCATCTGATTC
AGGCCAGTGAAAGACGTGTTACCTCGTCTGTCCCGCCAGGTTCCGCGAGCGGAGCCCTGACATCTTTCAGGAAG
CCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACTCCCAAGCCCCCTCCATCCTA
CAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCGGCATGACCGTCGCAGGGG
GAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTTGAGCCCGGAGGAGTCATAAGCAGAGATG
GAAGAATAAAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACTGACAGAGGTGAGCCGGAGTGAGGCAG
TGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGCCCCAGGAAG
ACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGAATGGTCCCCATCCTGGGTCA
TGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTCTGG
GCTTCTGCATTGTAGGAGGTTATGAAGAATAAATGGAAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGTGCTCAATGGTAGAAGTACATCAGGAA
TGATACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTG
GCACTTTTTTATTAGAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAATAGGCTAAGAAGTTGAAACACT
ATATTTATCTTGTGAGTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAA
TGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAAATTAATACTACTAGTTTTTTTTTTCAGTGTGGAGGAT
TTCTCATTACTCTACAACATTGTTTATATTTTTTCTATTCAATAAAAAAGCCCTAAAACAATAAATGATTGATT
TGTATACCCCACTGAATTCAAGCTGATTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTAT
GGCCATTTTTTAATTTACAGCTAAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAAACAAGAAT
AAATATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATA
 PSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSEFK
 KINRALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSES
 LSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPC
 QVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQGLGIKLVRKVDEPGVFIFNV
 LDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQE
 AGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYV
 ISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQED
 CSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGN
 KPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:**Signal peptide:**

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
 467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAA
 GCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCCTT
 CAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGAC
 AGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCCTAAAGCAAA
 GGAGTTCATGGCTAATTTCCATAAGACCCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGA
 AGCATCCACGAAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCA
 GAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATCCCAAAGT
 GTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTCGTTCC
 CCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTTCCTGCAGAG
 GCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGTTTAATCGAGC
 CAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGACTGCTTTATATT
 CCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAGCATCCCAA
 GCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGGGGG
 TGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGG
 ATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCG
 GCCCCTGCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGT
 GAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCACGAGTCTGGAGAACAGATGGGTTGAG
 TAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGA
 TTTCTGGTTTGGTGCATTGACCCTGGATCTTTTGGTGATGTTTGGAGAAGTGAATCTTTGTTT
 GCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTAAGAACCTGTTACAGCTCATT
 GTTGAGCTGAATTTTTCCTTTTGTATTTTCTTAGCAGAGCTCCTGGTGATGTAGAGTATAAA
 ACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATGAGGGTTAAATATTGTAATATGG
 ATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAAAATGAACGCTATTTGAGGACT
 CTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGATAAAAGGCCACAGGAAA
 TAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAAAGGTACGAAGA
 TACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGTGAGAAG
 GCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTGGGAATGAAGAG
 GTAGCAGGAGGGTGGAGTGTCGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGA
 TAGCCTTCAGGGGAGGACCTGCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTC
 TCTATTAGTTTTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGT
 TTACAAGTTTACATATTAATAATAATAATATGTCTATCAAATACCTCTGTAGTAAAATGTG
 AAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTL
TNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAI
LVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDC
FIFHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNN
YWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTD
GLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACG**ATGG**CACTGAGCTCCCAGATCTG
GGCCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACACA
ACAGACGGGACAACCTTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGAT
GCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGG
CTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACG**TAGA**ACCTACCTGCCCTGCCCCC
GTCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTG
GTTCTTTTGTTCCTTCCAAA
AAA

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCAC
CCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGT
CGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTAT
TGTCTGGACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGGCACTATCATAGT
GACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAG
CAAAGTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCA
GCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCAT
GGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGG
GGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGG
GTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAG
GAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGC
TGATGACCCAGATTCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCT
CTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGA
GAAGAAGAGAGTGGACATTTGTGCGGGAACTCCTAACATATGCCCCCATTTCTGGAGAGAACAC
AGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGATCCAGCAAATACGGT
TTACTCCACTGTGGAAATACCGAAAAAGATGGAAATCCCCACTCACTGCTCACGATGCCAGA
CACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCT
GCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTI
QPEGGTIIIVTQNRNRERVDFFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYE
HLSKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGES
DMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKR
ERQEEYIEEKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMEN
PHSLLTMPDTPRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAAT
GGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTC
AGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTCCCA
GGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAA
AGAGCGTGCTGCAACAACAGAAGTGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTC
ATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGT
AATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCAT
CCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAAT
AAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCT
GAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTTGGAATT
CTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTC
TCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGA
AAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAI
PATMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANC
EFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHLIHF
SVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGT
GAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGA
GCAGATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAA
CTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGC
CAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAG
CGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATG
CAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACATCAC
TCCAGGCTCTGCCACTACTTGCTTGAGACACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAAC
ACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGG
GCAGGCTGTTAGGCCCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCGCCCCACCACC
CCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKY
KSSQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAG
 GGAGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGC
 AAGGAGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAG
 TTCCAGGGGGCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCC
 CTGCAGAGCGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGA
 GACGCCCTGAGCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGGGCAGCTGGCTCT
 AAAGTCAGTGAGGGCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTT
 CCAGGCTTTGGCGCAGCAGATGCTTTGGGCAACAGGGTCGGGGAAGCAGCCCATGCTCTGGGA
 AACACTGGGCACGAGATTGGCAGACAGGCAGAAGATGTCATTTCGACACGGAGCAGATGCTGTC
 CGCGGCTCCTGGCAGGGGGTGCCTGGCCACAGTGGTGCTTGGGAAACTTCTGGAGGCCATGGC
 ATCTTTGGCTCTCAAGGTGGCCTTGGAGGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACT
 CCGTGGGTCCACGGATACCCCGGAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGGAGCT
 CCCTGGGGTCAAGGAGGCAATGGAGGGGCCACCAAACCTTGGGACCAACACTCAGGGAGCTGTG
 GCCCAGCCTGGCTATGGTTCAGTGAGAGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCCA
 CCATCTGGCTCAGGTGGAGGCTCCAGCAACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGC
 AGTGGCAGTGGCAGCAATGGTGACAACAACAATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGC
 AGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAGTGGCGGCAGCAGTGGTGGCAGCAGTGGCAAC
 AGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGTCCTCCTGGGGATCCAGCACCGGCTCCTCC
 TCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGACATAAACCCGGGTGTGAAAAGCCAGGG
 AATGAAGCCCGCGGGAGCGGGGAATCTGGGATTACAGGGCTTCAGAGGACAGGGAGTTTCCAGC
 AACATGAGGGAAATAAGCAAAGAGGGCAATCGCCTCCTTGGAGGCTCTGGAGACAATTATCGG
 GGGCAAGGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTTGGTGGAGTCAATACTGTGAAC
 TCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAAGAATTTTAAATCCAAGCTG
 GGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTCGCATCCCGTGTGACCTCCA
 GACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAAACACCACCCTCTCA
 TCACTAATCTCAGCCCTTGCCCTTGAAATAAACCTTAGCTGCCCCACAAAAAAAAAAAAAAAAA
 AA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCT
GGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCT
GGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGT
GTTTCCCACAGCCCCCAAACGGAAGTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAG
AGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGC
TGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATACCAATG
CCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGGCTGGGAG
AAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCT
TCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTG
ACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCC
TCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGC
CCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATA
TCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCT
GCCGCCTGGTGCATGACTTCACAGACGCTGTATCCGGGAGCGGCGTCGCACCCTCCCCACTC
AGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGGATTTTATTGATGTGC
TTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGATATAAGAGCAGAGGCTG
ACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTCCTGGGTCTGTACAACC
TTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGCAAGAGCTTCTGAAGGACC
GCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCCTTCTGACCATGTGCGTGA
AGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATCTCCCGATGCTGCACCCAGGACATTG
TTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCCTCATCGATATTATAGGGGTCC
ATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACCCCTTCCGCTTTGACCCAGAGA
ACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCTTTCTCCGAGGGCCCAGGAAGTGCATCG
GGCAGGCGTTTCGCCATGGCGGAGATGAAAGTGGTCCTGGCGTTGATGCTGCTGCACTTCCGGT
TCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGCGCGCCGAGGGCGGGC
TTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCATGACTTTCTGACCCATCCACCTGTT
TTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWGH
LGLITPTEEGLKDSTQMSATYSQGFTVWLGPPIPFIVLCHPDTIRSITNASAAIAPKDNLFIR
FLKPWLGE GILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRL
DMFEHISLMTLDSLQKCI FSDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHD
GRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFFKDKAKSKTLDFIDVLLLSKDEDGKALSD
EDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQL
PFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYD
PFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLEL
IMRAEGGLWLRVEPLNVGLQ

Important features:**Transmembrane domains:**

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPF
ARDAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAACATGCTCTGTCTGTGC
 CTGTACGTGCCGGTCATCGGGGAAGCCAGACCGAGTTCCAGTACTTTGAGTCGAAGGGGCTCCCTGCCGAGCTG
 AAGTCCATTTTCAAGCTCAGTGTCTTTCATCCCTCCCAGGAATTTCCACCTACCGCCAGTGGAAAGCAGAAAATT
 GTACAAGCTGGAGATAAGGACCTTGATGGGCGAGTACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAG
 AAGAAGCTGAGGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGCAG
 TCCCTGCCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAGCATGGATAAAAAACGGC
 ACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCCACCCCGTGGAAAACATCCCCGAGATCATC
 CTCTACTGGAAGCATTCACGATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTTACAGTGGAGGAG
 AGGCAGACGGGGATGTGGTGGAGACACCTGGCTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCC
 CCCCTGGACAGGCTCAAGGTGCTCATGCAGGTCCATGCCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTC
 ACTCAGATGATTTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAACGTCCAAAATTGCCCC
 GAATCAGCCATCAAAATTCATGGCCTATGAGCAGATCAAGCGCCTTGTGGTAGTGACCAGGAGACTCTGAGGATT
 CACGAGAGGCTTGTGGCAGGGTCTTGGCAGGGGCCATCGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAG
 ACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAG
 GGGTGGCCGCCCTTCTACAAAGGCTATGTCCCAACATGCTGGGCATCATCCCCTATGCCGGCATCGACCTTGCA
 GTCTACGAGACGCTCAAGAATGCCCTGGCTGAGCACTATGCAGTGAACAGCGCGGACCCCGGCGTGTGTGTCTC
 CTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGGCCAGCTACCCCTGGCCCTAGTCAGGACCCGGATG
 CAGGCGCAAGCCTCTATTGAGGGCGCTCCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAG
 GGGGCTTCCGGCTGTACAGGGGGCTGGCCCCAACTTCATGAAGGTTCATCCAGCTGTGAGCATCAGCTACGTG
 GTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTTCGGGTGACGGGGGAGGGGCCCGCCGAGTGGACTCG
 CTGATCCTGGGCCGAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACTAAGCTGTCTCGAGCC
 AAGCTGTGAAAACCTAGACGCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTTGTCTGTCTGACC
 CCAGCAGACCCTCCTGTTGGTTCCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCCAGGCTCCGGGCT
 CACATGTGTAAGGACAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGTTC
 TTCCATTTACCCCTTGCAGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGCCCT
 CTGCTGCCTGCCTGTCTGCTGAGGTAAAGGTGGGAGGAGGGCTACAGCCACATCCACCCCTCGTCCAATCCC
 ATAATCCATGATGAAAGGTGAGGTACGTGGCTCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAACTT
 GGCTGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCG
 GGCATGCTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCATGGTG
 CTCTGAGCTGGCCTGGACCCTGTGAGGATGGGCCCCACCTCAGAACCAAACTCACTGTCCCACTGTGGCATGAG
 GGCAGTGGAGCACCATGTTTGAAGGGCGAAGGGCAGAGCGTTTGTGTGTTCTGGGGAGGGAAGGAAAGGTGTTGG
 AGGCCTTAATTATGAGTGTGGGAAAAGGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGC
 TTCCAGAGGAAGACGAGGGAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCTGTG
 CAACCCAGCAGGGGCGCAGCGGGACCAGCCCCACATCCACTTGTGTCACTGCTTGGAACTATTTATTTTGT
 TTTATTTGAACAGAGTTATGTCCTAACTATTTTATAGATTTGTTTAAATTAATAGCTTGTCAATTTCAAGTTCAT
 TTTTATTCATATTTATGTTTCATGGTTGATTGTACCTTCCCAAGCCCGCCAGTGGGATGGGAGGAGGAGAGAA
 GGGGGGCTTGGGCCGCTGCAGTCACATCTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAG
 AAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAGGTTGGGGAAGGGCTTGCCCCAGCCTTAGGATTTAGGGTTT
 GACTGGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGA
 GGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAAGTGCCTCTCACTGTGAATTTGTGGTGGGCGG
 GGCTGGAGGAGAGGGTGGGGGGCTGGCTCCGTCCCTCCCAGCCTTCTGCTGCCCTTGTCTTAACAATGCCGGCCAA
 CTGGCGACCTCACGGTTGCATTCCATTCCACAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAGATC
 AATGCAAAAATTGTTATATATGAACATATAACTGGAGTCGTCAAAAAGCAAATTAAGAAAGAATTGGACGTTAGA
 AGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 58

MLCLCLYVPVIGEAFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDL
DGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILKSM
DKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRH
LVAGGGAGAVSRTCTAPLDRLKVLMMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLK
IAPESAIFMAYEQIKRLVGSQDQETLRIHERLVAGSLAGAIQAQSSIYPMEVLKTRMALRKTGQ
YSGMLDCARRILAREGVAAFYKGYVPNMLGIIIPYAGIDLAVYETLKNWLQHYAVNSADPGVF
VLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFLYRGLAP
NFMKVIPAVSISYVYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATG
GCTTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCA
ATTGCACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCC
TCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCT
TCTGATATCGTGATAACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAA
GGCAAAGATGAGCTGTGCGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTGCTGAT
CAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACC
TACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGA
GCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAG
GCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAAC
TTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTGTG
TCTGTGCTCTACAATGTTACGATCAACAACATACTCCTGTATGATTGAAAATGACATTGCC
AAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAGGCCGAGTCACCTACAGCTG
CTAAACTCAAAGGCTTCTCTGTGTGTCTCTTCTTTGCCATCAGCTGGGCACTTCTGCCT
CTCAGCCCTTACCTGATGCTAAAATAATGTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTG
TTACAACAGGGATCTACAGAACTATTTACCACCAGATATGACCTAGTTTTATATTTCTGGGA
GGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCAAGAAACAAAAAGAAGCC
AAAAGCAGAAGGCTCCAATATGAACAAGATAAATCTATCTTCAAAGACATATTAGAAGTTGGG
AAAATAATTCATGTGAAGTAGACAAGTGTGTTAAGAGTGATAAGTAAATGCACGTGGAGACA
AGTGATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGGAGTGAGAGGACAGGAT
AGTGATGTTCTTTGTCTCTGAATTTTATAGTTATATGTGCTGTAATGTTGCTCTGAGGAAGCC
CCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCACAAATTAAGCTGTAGTATGTAC
CCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCGGCTGCATTTTAGTAATGGG
TCAAATGATTCATTTTTATGATGCTTCCAAAGGTGCCTTGGCTTCTCTTCCCAACTGACAAA
TGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGTCGGGGACACCGATTT
TATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIK
LSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAG
TYKCYIIITSKGKGNNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGA
NFSEVSNTSFELNSENVTMKVVSVLNVTINNTYSCMIENDIAKATGDIKVTSESEIKRRSHLQ
LLNSKASLCVSSFFAISWALLPLSPYLMLK

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACA
AGCACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGG
TAGTGGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACC
AGCTGGTGGAGGGCCCTATGGACACCCAATCCTGGGATGTTCCCCTCTGGAActCCAGGAGG
ACCATATGGCGGTGCAGCTCCCGGGGGCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGC
CCAGCAGCCTGGGCTTTATGGACAGGGTGGCGCCCCTCCCAATGTGGATCCTGAGGCCTACTC
CTGGTTCCAGTCGGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGC
CCTGGTCAACTGCAATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTT
TGACAAGACCAAGTCAGGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCA
GCAGTGAAGAACCTCTTCCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGA
GCTGCAGCAAGCTCTGTCCCAAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTTCTGGT
CTCCCGCTACTGCCCACGCTCTGCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTG
CACCCAGCTGCAGGTGCTGACAGAGGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACAT
CCGGCTCAGCTTCGAGGACTTCGTCAACATGACAGCTTCTCGGATGCTATGACCCAACCATCT
GTGGAGAGTGGAGTGCACCAGGGACCTTTCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACA
TCTCTTCTTTTCTGTCCCTCTAGAAGAACATTCTCCCTTGCTTGATGCAACACTGTTCCAAA
AGAGGGTGGAGAGTCCTGCATCATAGCCACCAAATAGTGAGGACCGGGGCTGAGGCCACACAG
ATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGAATGTCCTGATGGCCATGAGCAGTTGAGTG
GCACAGCCTGGCACCAGGAGCAGGTCTTGTAATGGAGTTAGTGTCCAGTCAGCTGAGCTCCA
CCCTGATGCCAGTGGTGAAGTGTTCATCGGCCTGTTACCGTTAGTACCTGTGTTCCCTCACCAG
GCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGAATCTGACCAAGCATGAGAGAGATCT
GTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAATCCTTGTGTGTTAACTTCTAGC
TGCTTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCATCTTTGGCCAGGCTTCTGCC
CCCTGCAGCTGGGACCCCTCACTTGCTGCCATGCTCTGCTCGGCTTCAGTCTCCAGGAGACA
GTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTTCATTTGGGGCCAAAAG
TCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLP PGGGYGGPAPGGPYGPPAGGGP
YGHNPNGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSV
DSDHSGYISMKELKQALVNCNWSSFNDCTCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNL
FQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQV
LTEAFREKDTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCA
TCTCCCAGGCCTCTTTGCCCCGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGG
GACCAACTTGCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCA
GCCCCGCTCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCCTCTGAAGCTCAGCGTGCCTCC
ATCAGATGGCTTCCCACCTGCAGGAGGTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCT
GCCTGCCATGGATTCTTGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGGCTGAGGA
CCGCCTGGGGGAAGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGGCCCTCGCTCC
GGGCAGTGGCCCTTTGCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTC
ACTCCTCCACCAGGACTCGGAGTCCAGACGACTGCCCCGTTCTAATTCAGTGGGAGCCGGGGG
AAAAATCCTTTCCCAACGCCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTG
GGGTACCCTGAATCCCAGTGTGTCTGCGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAG
GCCCATGCCACACCCTGAGGGAATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGG
AAATATTAATCGGTATCCAGGAGGCAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTG
GGGGAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTCATCTATACCCAGGTATCAA
TAACCCATTTCTCCTGGAGTTCTCCGCCCTCCTGGCTCTTCTTGGAACATCCCAGCTGGCTT
CCCTAATCCTCCAAGCCCTAGGTTGCAGTGGGGCTAGAGCACGATAGAGGGAAACCCAACATT
GGGAGTTAGAGTCCTGCTCCCGCCCCCTTGCTGTGTGGGCTCAATCCAGGCCCTGTTAACATGT
TTCCAGCACTATCCCCACTTTTCAGTGCCTCCCCTGCTCATCTCCAATAAAATAAAAGCACTT
ATGAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQP
ALDPRSNDLARVPLKLSVPPSDGFPPAGGSQVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDR
LGEALPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGK
ILSQRPPWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWN
INRYPGGSWGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNI PAGFP
NPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATG**GGG
TCTGGGCTGCCCCCTTGTCCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGT
ATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTC
CTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCCTC
CACCATGCAAGATCTCAACACCATGTTGTCTGCAACACAT**TGA**CAGCCATTGAAGCCTGTGTCC
TTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCGACCCTGTCTTTCAGCAG
GCCCCCACCCTCCTGAGTGGCAATAAATAAAATTTCGGTATGCTG

FIGURE 66

MGSGLPVLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSEYESSFLELLEKLCLLLHLPSTSV
TLHHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGG
GCCAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAG
CCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTT
CTGGCGCTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACC
ACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAAC
CTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTG
GCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTAC
CGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAG
GAGACGGTGCAGGGCTGCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGC
TGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAG
CAGGGAAGAAGGTACTTCAAAGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTT
TTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIVV
FSLLAALLLAVGLALLVRKLRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYL
 FEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGE
 KGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCSA
 GISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNE
 KTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPVFSLLKISQRIVCLVLDKSG
 SMGGKDRLNRMNQAAKHFLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTY
 PLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTGDGENTASSCIDEVKQSGAIVHFIALGRA
 ADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMN
 DTVIIDSTVGKDTFFLITWNSLPPSISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAY
 NLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFFSPMIVYAEILQGYVPVLGANVT
 AFIESQNGHTEVLELLDNGAGADSFKNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRP
 PLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPS
 QITDLDATVHEDKIIILTWTAPGDNFVGVQRYIIRISASILDLRDSFDDALQVNTTDLSPKE
 ANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSANIAQVTLFIPQANPDDIDPTPTPT
 PTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCGGGCAGGGG
 TGACAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCCTATTTCAAGGAAAGACGCCAAGGTAATTTT
 GACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCCAGTTATGCCAGGATTTACTAG
 AGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCTTTGTGGGGCTGC
 GTTCTCTTAGCAGTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCCCTCTTGCTGTTG
 GCTGCACATCAGGAAGGCTGTGATGGGAATGAAGGTGAAAACCTGGAGATTTCACTTCAGTCATTGCTTCTGCCT
 GCAAGATCATCCTTTAAAAGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAA
 GGAAATGGATGCAAGCAGCTCCGGGGGGCCCCAACGCATGCTTCCTGTGGTCTAGCCAGGGAAAGCCCTTCCGTG
 GGGCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCCTGAATGATGATTGCTGCGCGGGGG
 CTGCTTGCCTGGATTTCCTGGGTGGTGGTTTCTGGTGGTCTCTCTGCTGTGCTATCTCTGCTGTACATGTTG
 GCCTGCACCCCAAAAGGTGACGAGGAGCAGCTGGCAGCTGCCAGGGCCAACAGCCCCACGGGGGAAGGAGGGGTAC
 CAGGCCGTCCTTCAGGAGTGGGAGGAGCAGCACCGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTC
 AAGGAGGAGCTGCAGGAGAGGAGTGAGCAGCTCAGGAATGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGT
 CTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAG
 GTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTGCTTTTCGATAGCTTTACTCTACAGAAGGTGTAC
 CAGCTGGAGACTTTCCCGGCTTGGAGACCTGAACAATCCTGCAGAGAACAGCCCAATCACCCTCCTTACACGGCCTCT
 GCCATTGAATCAGCCTTGGAGACCTGAACAATCCTGCAGAGAACAGCCCAATCACCCTCCTTACACGGCCTCT
 GATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCCTCAAAGGGGACCAC
 AAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTCAGCCCCATCATGAAAGTGAAAAATGAAAAGCTCAAC
 ATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTCGGGCAGTTTCATGCAGAAT
 TTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGGAAAGAAGAAATAAAT
 GAAGTCAAAGGAATACTTGAACACACTTCCAAAGCTGCCAACTTCAGGAACCTTTACCTTCATCCAGCTGAATGGA
 GAATTTTCTCGGGGAAAGGACTTGATGTTGGAGCCGCTTCTGGAAGGGAAGCAACGTCCTTCTCTTTTCTGT
 GATGTGGACATCTACTTCACATCTGAATTCCTCAATCCTGCTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTT
 TATCCAGTTCTTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAG
 CAGCTGGTCATAAAGAAGGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTCAGTATCGGTGAGAC
 TTCATCAATATAGGTGGGTTTTGATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTAT
 CTCCACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATG
 GACGAGCTGACCCCGAGCAGTACAAGATGTGCATGCGAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTG
 GGCATGCTGGTGTTCAGGCACGAGATAGAGCTCACCCTCGCAACAGAAACAGAAGACAAGTAGCAAAAAACA
TGAACTCCAGAGAAGGATTGTGGGAGACACTTTTTCTTTCTTTTGAATTACTGAAAGTGGCTGCAACAGAGA
 AAAGACTTCCATAAAGGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGT
 TGGGCTTTTTTACAACAGAAATCAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCT
 GACAAAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTTACAATACACTGAGA
 CCTGTTGTTTTGTGTGCTCATTGAAATATTCATGATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGG
 CAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGG
 CAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAGTAAAATAAAATGGACCAGAAAAGAAAAGAA
 ACCATAAATATCGTGTGCATATTTTCCCAAGATTAACCAAAAATAATCTGCTTATCTTTTGGTTGCTCTTTTAA
 CTGTCTCCGTTTTTTCTTTTATTTAAAAATGCACTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTAC
 CACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTACATTTTTTATATTTTTTAAGAAGATACTTTGAGATGCA
 TTATGAGAACTTTCAGTTCAAAGCATCAAATTGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTCAGGC
 ACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGACGTACAGATACTTTCTCTGAAG
 AGTATTTTCAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAAAAGGAACT
 CATTACAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCAATCTCTTTTCAAACAGGGGTGCTCCTCCTGG
 CTTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATTGTGAAAGATCAATCCATCTG
 CCAGAATCTAGTGGGATGGAAGTTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAGTAACTGAATTATTTT
 TTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTTCAAACATTTTTT
 TAAAAATAAATACAGTTAAACATAGAGTGGTTTCTTCATTCATGTGAAAATATTAGCCAGCACCAGATGCATGAG
 CTAATTTCTCTTTGAGTCTTGTCTTCTGTTTGTCTCAGATAAATCATTGTTTTAAAGCTTCAAGAACATTCAA
 GCTGTTGGTGTGTTAAAAATGCATTTGATTGATTGCTGAGTTTATGAAATTTAATTTAAACACAGGCCA
 TGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTTGTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQE
WEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHS
QVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALE
TLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVK
NEKLNMAANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILEN
TSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPG
KKVFYPVLFSQYNPGIIYGHHDVPPLEQQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDL
DIKGWGGEDVHLYRKYLHSNLIVVRTPVRLFWHEKRCMDLTPEQYKMCMQSKAMNEASH
GQLGMLVFRHEIEAHLRKQKQKTSSKKT

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTTGTCCTGGGGATC
CAGAAACCCATGATACCTACTGAACACCGAATCCCCCTGGAAGCCCACAGAGACAGAGACAGC
AAGAGAAGCAGAGATAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCTCTCTCA
CTCCTCCCTCCCTCTCTCTCTGCTGTCTAGTCCTCTAGTCCTCAAATTCCCAGTCCCCTGC
ACCCCTTCCTGGGACACT**ATG**TTGTTCTCCGCCCTCCTGCTGGAGGTGATTTGGATCCTGGCT
GCAGATGGGGGTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGGCCAGCCTCT
TACCCTGAGTGTGGAACAATGCCAGTCGCCATCGATATTCAGACAGACAGTGTGACATTT
GACCCTGATTTGCCTGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTTGGAC
CTGCACAACAATGGCCACACAGTGCAACTCTCTGCCCCTCTACCCTGTATCTGGGTGGACTT
CCCCGAAAATATGTAGCTGCCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGG
TCAGAACACCAGATCAACAGTGAAGCCACATTTGCAGAGCTCCACATTGTACATTATGACTCT
GATTCCTATGACAGCTTGAGTGAGGCTGCTGAGAGGCCTCAGGGCCTGGCTGTCCTGGGCATC
CTAATTGAGGTGGGTGAGACTAAGAATATAGCTTATGAACACATTCTGAGTCACTTGCATGAA
GTCAGGCATAAAGATCAGAAGACCTCAGTGCCTCCCTTCAACCTAAGAGAGCTGCTCCCCAAA
CAGCTGGGGCAGTACTTCCGCTACAATGGCTCGCTCACAACCTCCCCCTTGCTACCAGAGTGTG
CTCTGGACAGTTTTTTATAGAAGGTCCCAGATTTCAATGGAACAGCTGGAAAAGCTTCAGGGG
ACATTGTTCTCCACAGAAGAGGAGCCCTCTAAGCTTCTGGTACAGAACTACCGAGCCCTTCAG
CCTCTCAATCAGCGCATGGTCTTTGCTTCTTTCATCCAAGCAGGATCCTCGTATACCACAGGT
GAAATGCTGAGTCTAGGTGTAGGAATCTTGGTTGGCTGTCTCTGCCTTCTCCTGGCTGTTTAT
TTCATTGCTAGAAAGATTCGGAAGAAGAGGCTGGAAAACCGAAAGAGTGTGGTCTTCACCTCA
GCACAAGCCACGACTGAGGCAT**TAA**ATTCCTTCTCAGATACCATGGATGTGGATGACTTCCCTT
CATGCCTATCAGGAAGCCTCTAAAATGGGGTGTAGGATCTGGCCAGAAACACTGTAGGAGTAG
TAAGCAGATGTCCTCCTTCCCCTGGACATCTCTTAGAGAGGAATGGACCCAGGCTGTCATTCC
AGGAAGAAGTGCAGAGCCTTCAGCCTCTCCAAACATGTAGGAGGAAATGAGGAAATCGCTGTG
TTGTTAATGCAGAGANCAAACTCTGTTTAGTTGCAGGGGAAGTTTGGGATATACCCCAAAGTC
CTCTACCCCTCACTTTTATGGCCCTTCCCTAGATATACTGCGGGATCTCTCCTTAGGATAA
AGAGTTGCTGTTGAAGTTGTATATTTTTGATCAATATATTTGGAAATTAAAGTTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPAASYPECGNNAQSPIDIQTDSVTFDPDLPA
LQPHGYDQPGTEPLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQIN
SEATFAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQ
KTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTE
EEPSKLLVQNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKI
RKKRLENRKSVVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTG
TCTCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGA**ATG**TCC
TACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTCA
TCACAAATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTG
ACAGCTCCAGAGAAGTGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATAC
TCCAATCTGAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGT
GTGACCAACCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTG
GAGTCCTTCGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACT
TTGAAAGATCAATCATCAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCT
ATTACCGTGTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAA
GAGAAACACCCAGCAAATTTGATTTTGAATTTATGGAAATGAATTTGACAAAAGATTCTTTGTG
CCTGCTGAAAAAATCGTGATTAACCTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCT
CATCAGGATATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCC
AGCGGGAACCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCAT
TTGATGGAAATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTACCCAGCAAGAG
TCCCTCAGCAGAACAATACCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACC
ACTGACATTTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAA
GGAACATTATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCGCAAACGTTACAGTACTCA
TACACCCCTCAGCTCCAAGACTTAGACCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGG
CCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAACTGGCAGGCTGTGTATT
CCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTC
GGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAA
AATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGGTATATGTGCAGATGGAAAAC**TGA**
TGCCAACACTTCCTTTTGCCTTTTGTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCA
GCCATAAAGTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTGAGTGTCTGTGAGAATTACTT
ATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTAGGTCTCTTAACAA
TGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAAAGCAGTCAATA
AATGTTTGCCAGACTGGGTGCAGAATTTATTACAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSISQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQ
IYSNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCA
RTLKDQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFEDKRF
FVPAEKIVINFITLNISSDKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYA
SHLMEIFCDSEENTEGTSLTQQESLSRTIPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVS
TQGTLLSQAALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRL
CIPSLSSFDQDSEGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCAGGTCTGGCATCCTG¹CACTTGCTGCCCTCTGAC
ACCTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTT
GATCCAAGCCACCCTCAGTCCCAGTGCAGTTCTCATCCTCGGCCCAAAGTCATCAAAGAAAA
GCTGACACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAG
TGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCT
GAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTC
GGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCC
CCTGGTCAAGACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGA
CACCAGTGCAAGTGGCCCCACCCGCCTGGTCCTCAGTGA²CTGTGCCACCAGCCATGGGAGCCT
GCGCATCCA³ACTGCTGTATAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGCAGGTCATGAA
CCTCCTAGTGCCATCCCTGCCCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTC
CTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGA
CCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGG
GGCCAAGTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCTT
GACAATGCCCCACCCTGGACAACATCCCGTTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGC
TGCAGTGGCTGCTGTGCTCTCTCAGAAGAATTATGGTCCTGTTGGACTCTGTGCTTCCTGA
GAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGAAAAGGCTGCAGATAAGCTGGG
ATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTTTTTTATAGACCAAGGCCA
TGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTTCCCTCCAGTGAAGCCCTCCGCCCTTT
GTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTACACCAAAGGTGACCAACTTAT
ACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTGATGAACTCTGGGATTGGCTGGTT
CCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCACTCCATCCTGCTGCCGAACCA
GAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGGCCTTGGGATTGAGGCAGC
TGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCCTTGTGGAAACCCAGCTC
TCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGGGTCCCAGCTGGGAGT
ATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTGCCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIEKLTQELKDHNATSILQQLPLLSAMR
 EKPAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVK
 TIVEFHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLV
 PSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKL
 LDSQGKVTKWFNNSAASLTMPITLDNIPFSLIVSQDVVKAAVAVALSPEEFMVLLDSVLPESAH
 RLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTL
 GIEASSEAQFYTKGDQLILNLNLISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGK
 LRSGVPVSLVKALGFEEAESSLTKDALVLT PASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGA
GCTTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGC
CTCTCTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGT
TGCCATGCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGT
TGGCTTCTCCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGA
CATCTATAGCACCCCTTCTGGGCCTGCCCCGCTGACATCCAGGCTGCCCAGGCCATGATGGTGAC
ATCCAGTGCAATCTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTT
CTGCCAGGAATCCCGAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGG
AGGCCTCCTGGGATTCATTCCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTC
ACCACTGGTGCCTGACAGCATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTTC
TTCCCTGTTCTCCCTGATAGCTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCG
CTCCAATACTACGATGCCTACCAAGCCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGG
TCAACCTCCCAAAGTCAAGAGTGAGTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAGAAC
CAGGGGCCAGAGCTGGGGGGTGGCTGGGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCA
CAGGTGAGGGACACTACCACTGGATCGTGTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGG
CCATTGGATTGAGCAAAGGCAGAAATGGGGGCTAGTGTAACAGCATGCAGGTTGAATTGCCAA
GGATGCTCGCCATGCCAGCCTTTCTGTTTTCTCACCTTGCTGCTCCCCTGCCCTAAGTCCCC
AACCCTCAACTTGAAACCCCATTCCTTAAGCCAGGACTCAGAGGATCCCTTTGCCCTCTGGT
TTACCTGGGACTCCATCCCCAAACCCACTAATCACATCCCCTGACTGACCTCTGTGATCAA
AGACCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATTGCTGGGGATGGGAAGGAGAAGCAGT
GGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTCCCTCCAAAGAACTGATTGGCCC
TGGAACCTCCATCCCCTCTTGTATGACTCCACAGTGTCCAGACTAATTTGTGCATGAACTG
AAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATGCAGGATGGGAGGACAGGAA
GGCAGCCTGGGACATTTAAAAAATA

FIGURE 80

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQ
CDIYSTLLGLPADIQAAQAMMVTSSAIISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFI
LGGLLGFI PVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQR
NRSNYDAYQAQPLATRSSPRPGQPPKVKSEFNSSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTT
TCCCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCCTT
TCCTGCCCCACCGCTGCTTCCTGGCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGT
CTGTGGGTTGATCTGTGGCCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGACTCC
GCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCTCTCCTC
CTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACAT
GTTCTGCCTTTTCCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCC
ACAAGGCCTGATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCG
CCTCCACTGTCCGCTGTCCACTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAA
GTGTGTGGAACCTCACACTCCCTCTGGACTCCGGGGCCCCACCAAAGTCCTGCCAGCACAACGG
GACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCCTGCCCAA
CCAGTGTGTCCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGA
ACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGC
AAGTGAGCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGA
TCCATGTTCCAGTGATGCTGGGAGAAAGAGAGGGCCGGGCACCCCAAGCCCCACTGGCCTCAG
CGCCCCCTCTGAGCTTCATCCCTCGCCACTTCAGACCCAAGGGAGCAGGCAGCACAACCTGTCAA
GATCGTCCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGGCGGGAAGACGTACTCCACGG
GGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTGCATCCTATGCACCTGTGA
GGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGTACCCCTGCCGTACCCCGA
GAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGGACAAAGCAGACCCTGGCCACAGTGA
GATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTCCTCGTCCACACATCGGTATCCCC
AAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGCCTCGGACTTGGTGGAGATCTA
CCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAGGTGAAGTACCTGGCCCAAG
GCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAGGAAGCAAGACTTCCAGA
AAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGGTCACTGGAACGTCTTCCTAG
CCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAGACATAACAAAGAC
CTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAATAAGAAGTTGC
ATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTC
SEGAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSA
HELFP SRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQ
SLHGVRHPQDPCSSDAGRKRGPPTPAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKAC
VHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKVKAGKCKICP
EDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETE
AQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSPDPGAEGHGQS
RQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGA
GCTTCTCCGTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCC
GTCACCTCTCCTGTCATCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTC
TCATGCTCAGTTTGGTTCTGAGTCTCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGC
CAGACAAGCCTGTCCAGGCCTTGGTGGGGGAGGACGCAGCATTCTCCTGTTTCTGTCTCCTA
AGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGGCCAGTTCTCTAGCGTGGTCCACC
TCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATCAAGGCAGGACAAAAC
TGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACTGTGTTGG
ATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGGAGC
TACAGGTGTCAGCACTGGGCTCAGTTCCCTCTCATTTCCATCACGGGATATGTTGATAGAGACA
TCCAGCTACTCTGTGAGTCCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCAC
AAGGACAGGATTTGTCCACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGG
AGATCTCTCTGACCGTCCAAGAGAACGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATC
TGAGCCGAGAGGTGGAATCCAGGGTACAGATAGGAGATACCTTTTTTCGAGCCTATATCGTGGC
ACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTATTTTTTGGCATTGTTGGACTGA
AGATTTTCTTCTCCAAATTCAGTGGAATCCAGGCGGAACTGGACTGGAGAAGAAAGCACG
GACAGGCAGAATTGAGAGACGCCCCGGAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGG
CTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATAGAAAAGCTCCCCAGGAGG
TGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAG
GGAAACATTACTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGG
ATGATGTGGACAGGAGGAAGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCTCA
GACTGAATGGAGAACATTTGTATTTACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGA
CCCCACCTACAAAAATAGGGGTCTTCTGGACTATGAGTGTGGGACCATCTCCTTCTTCAACA
TAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTGAAGGCTTATTGAGGCCCTACA
TTGAGTATCCGTCCTATAATGAGCAAAATGGAATCCCATAGTCATCTGCCCAGTCACCCAGG
AATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAGTGAGT
CCTCCTCACAGGCAACCACGCCCTTCTCCCCAGGGTGAAATGTAGGATGAATCACATCCCA
CATTCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAA
GGTGGCTTCCAGATGAAGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCT
GAGCTGGGAGGGAAGAAGGCTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGA
TCTTGAAATACCACCTCTCAGGTGAAGAACCGTCAGGAATTCCCATCTCACAGGCTGTGGTGT
AGATTAAGTAGACAAGGAATGTGAATAATGCTTAGATCTTATTGATGACAGAGTGTATCCTAA
TGTTTGTTCATTATATTACACTTTCAGTAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSQWQVFGPDKPVQALVGEDAAFSCFLSPKTNAEAMEVRFFRGQFSSV
VHLYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAI
WELQVSALGSVPLISITGYVDRDIQLLCQSSGWFFRPTAKWKGPQGQDLSTDSRTNRDMHGLF
DVEISLTVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIV
GLKIFFSKFQWKIQAELDWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAP
QEVPHSEKRFTRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYW
VLRNLNGEHLFTLNPRFISVFPRTPTTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLR
PYIEYPSYNEQNGTPIVICPVTQESEKEASWQRASAIPESTNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCAGAC**ATG**CTGCTGCTGCTGCTGCCC
CTGCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCC
GTGACGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGC
TGGATTTACCCTGGCCCAGTAGTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAG
GATGCTCCAGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTC
CACCTCCTTGGGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGT
GATGCGGGGAGATACTTCTTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCAC
CGGCTCTCTGTGAATGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTG
GAGTCCGGCTGCCCCCAGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCC
CCTATGATCTCCTGGATAGGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCG
GTGCTCACCCCTCATCCCACAGCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTC
CCTGGGGCCAGCGTGACCACGAACAAGACCGTCCATCTCAACGTGTCCTACCCGCCTCAGAAC
TTGACCATGACTGTCTTCCAAGGAGACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCT
CTGTCACTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAAT
CCCCCTGCCAGGCTGAGCCTGAGCTGGAGAGGCCTGACCCTGTGCCCCCTCACAGCCCTCAAAC
CCGGGGGTGCTGGAGCTGCCTTGGGTGCACCTGAGGGATGCAGCTGAATTCACCTGCAGAGCT
CAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCCCTGCAGAGCAAAGCCACATCA
GGAGTGA CT CAGGGGGTGGTTCGGGGGAGCTGGAGCCACAGCCCTGGTCTTCTCCTGTCCTTCTGC
GTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAATCGGCAAGGCCAGCAGCGGGCGTGGGA
GATACGGGCATAGAGGATGCAAACGCTGTCAGGGGTTACGCCTCTCAGGGGCCCCCTGACTGAA
CCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAGCTTCTGCCCCTCCTCAGTGGGG
GAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTGAAGCCTTGGGACTCGCGGGGA
CAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAGAT**TGA**GAAACTGCAGAGACT
CACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCAGAGGCTGATTCTTGTAG
AATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATGTGCAGAGTGAAAAGC
ACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCCTCCCTTTTATTTT
TTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWFR
EGANTDQDAPVATNNPARAVWEETRDREHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGS I
KWN YKHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPLD
PSTRSSVLTLPQPQDHGTS LTCQVT FPGASVTTNKT VHLNVSYPQNLTM TVFQGDGT VST
VLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDA
AEFTCRAQNPLGSQQVYLNVSLSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRK KSA
RPAAGVGDTGIEDANAVRG SASQG PLTEPWAEDSPPDQPPPASARSSVGE GELQYASLSFQMV
KPWDSRGQEATDTEYSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGT
ACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTG
TTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGG
ACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGT
GCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGAC
ATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGG
AAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGG
GACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTAC
AAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAG
TCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTC
CAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGA
AAGTGTTGGA CTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAA
ACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGG
GTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAAC
ACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGA
GATTTTCTGGTTTTTGATTGGAGTGGATATGGAAC TCATGTTGGTTACAGCAGCAGCCGTGAG
ATAACTGAGGCAGCTGTGCTTCTATTCTATCGTTGAGAGTTTTGTGGGAGGGAACCCAGACCT
CTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAAC TTACCCAGTAGCTAGAATGTTAATG
GCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

FIGURE 88

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENG
VIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSA
EAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGLIYQK
YPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYSPYGQREFTAGFVQFRVFNNERAANALCA
GMRVTGCNTEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:**Signal peptide:**

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAG
ACCGCCGCCCTTGTCCCCGAGGGGCC**ATG**GGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCT
GACGCTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACA
GGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGC
CGCGCTCTCTGTACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTC
CATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCT
GTCTTCTTCATATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTGTCTTCTGCAG
TGCCCTTCCAGCTGTCACTGAAATGGCTTTATTCGTCACCGTCTTTGGGCTGAAAAGAAACC
CTTCT**TGA**TTACCTTCATGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGT
ATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTCTGCTGGAGGA
TATGTGTTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAA
TAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAA
AAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGL
FAVELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEM
ALFVTVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTG
 ACTCGCTGCTGCTTCGTGTTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGAT
 GATGGTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGGCCAGCCCCGGCCCCGG
 GTGCCTCGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTA
 GGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCCGAAC
 CACAGCCCCCACCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAAC
 ATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACC
 TTCAGCGTCCACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGGCC
 CCCAGTAAAGCTGTAGAGTTCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAA
 ATCTTCAACTGCCGGATGGAGTGGGAGAAGGTAGAACGGGGCCCGCGGACCTCGCTTTGCACC
 CACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCC
 CAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAG
 AAGGTGTGCCCAGATTACAACCTACCATAGTGATACCCCTACTACCCATCTGGGTGAACCCGGG
 GCAGGCCACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGG
 ACACCGGGCAGGGAAGGGGTGGGCCTCAGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAA
 GTGGGGCCAGGGCCAAGTCTCAAGTGGCAGAGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAA
 GCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGAGTGGGCTCTCTGTGCAGCCTCACA
 GGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCCTGTGGGCAGGCCGATCAG
 TGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTTCTTGCCATCCTGAGGAAAG
 ATAGCAACAGGGAGGGGGAGATTTCATCAGTGTGGACAGCCTGTCAACTTAGGATGGATGGCT
 GAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGGGCCAGAGGAGCTCTCCAGCCCTG
 CCTAGTGGGCGCCCTGAGCCCCTTGTCGTGTGCTGAGCATGGCATGAGGCTGAAGTGGCAACC
 CTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGCCAGGCCACCCCTTTCCAAA
 ATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATGGCACACCCATCCTTAAG
 CTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCCACAGCCCATCCGCGTGCTGTGTGTC
 CCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCGGAGAGGGGTCCCT
 CAACAGTCAGCCTCACCTGTCAGACCGGGTTCTCCCGGATCTGGATGGCGCCGCCCTCTCAG
 CAGCGGGCACGGGTGGGGCGGGGCCGGGCCGAGAGCATGTGCTGGATCTGTTCTGTGTGTCT
 GTCTGTGGGTGGGGGGAGGGGAGGGAAGTCTTGTAACCGCTGATTGCTGACTTTTGTGTGA
 AGAATCGTGTTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANS
TLLGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHG
NGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTS
LCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCG
CCCTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTT
TCTTCTGGTTGGTGTCTCTACTGATTTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTG
ACAACAAAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATA
TCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGA
GTATAAACCCAGGTGAGACAGCACCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCT
TTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCA
CAGTGGGCATTCATGGAGATTCTCCTCAATTCTTCCTTTATTAGCTTTCATGACGCTGGTCA
TTATCTTGCTGCATGTATTCTGGGGCATTGTATTTTTTTGATGGCTGTGAGAAGAAAAAGTGGG
GCATCCTCCTTATCGTTCTCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTT
ATTATGGAATAAACCTGGCGTCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCT
TAGCTGCGGGAGGCAGCTGCCGAAGCCTGAACTCTGCCTGCTCTGCCAAGACAAGAACTTTC
TTCTTTACAACCAGCGCTCCAGATTAACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACA
TCTTTAGAGGAAGCACAACTGTGCCTTTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAA
GAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFI AFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWF MARVIIDNK
DGPTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGI
MSGVFSEFVNTLSDSLPGPTVGIHGDS PQFFLYSAFMTLVII LLHVFWGIVFFDGCEKKKWGIL
LIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGTWAF LAAGGSCRS LKLCLLCQDKNFLLY
NQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
21

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCT
 CTCTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCCTGAC
CATGGTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCCA
 GCCTGCAGAGCTGTCTGTGGAAGTTCCAGAAAACCTATGGTGGAAATTTCCCTTTATACCTGAC
 CAAGTTGCCGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAGGGGACTCAGGCAA
 GGCAACTGAGGGGCCATTTGCTATGGATCCAGATTCTGGCTTCCTGCTGGTGACCAGGGGCCCT
 GGACCGAGAGGAGCAGGCAGAGTACCAGCTACAGGTCAACCCTGGAGATGCAGGATGGACATGT
 CTTGTGGGGTCCACAGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTT
 CTCTCAAGCCATCTACAGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCCTCTT
 CCTTGAGGCTTCAGACCGGGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCT
 GAGCCAGGCTCCAGCCCAGCCTTCCCCAGACATGTTCCAGCTGGAGCCTCGGGCTGGGGGCTCT
 GGCCCTCAGCCCCAAGGGGAGCACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTT
 GGTACAGGTCAAGGACATGGGTGACCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGT
 CTCCATCATAGAGAGCACCTGGGTGTCCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGT
 CCTATACCCGCACCATGCCCCAGGTACACTGGAGTGGGGGTGATGTGCACTATCACCTGGA
 GAGCCATCCCCCGGGACCCCTTTGAAGTGAATGCAGAGGGAAACCTCTACGTGACCAGAGAGCT
 GGACAGAGAAGCCCAGGCTGAGTACCTGCTCCAGGTGCGGGCTCAGAATTCCCATGGCGAGGA
 CTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTGATGAGAATGACAACGTGCCTATCTG
 CCCCCTCGTGACCCACAGTCAGCATCCCTGAGCTCAGTCCACCAGGTACTGAAGTGACTAG
 ACTGTCAGCAGAGGATGCAGATGCCCCCGGCTCCCCCAATTCCCACGTTGTGTATCAGCTCCT
 GAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTCCAGGTGGACCCCACTTCAGGCAG
 TGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACATCCTGCTTCTGGTGCTGGCCAT
 GGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAGTCGAAGTCGCAGTCACAGA
 TATCAATGATCACGCCCCCTGAGTTCATCACTTCCAGATTGGGCCTATAAGCCTCCCTGAGGA
 TGTGGAGCCCCGGGACTCTGGTGCCATGCTAACAGCCATTGATGCTGACCTCGAGCCCGCCTT
 CCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTGGCCTGGATTGGGA
 GCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAGGCAGCTCCAAG
 TCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCCAGGCCCTGG
 AGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCCAAGTTGGACCAGGA
 GAGCTACGAGGCCAGTGTCCCATCAGTGCCCCAGCCGGCTCTTTCCTGCTGACCATCCAGCC
 CTCCGACCCCATCAGCCGAACCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCTCTG
 CATTGAGAAATTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGGA
 CACCTACACGGTGCTTGTGGAGGCCCAGGATACAGCCCTGACTCTTGCCCCCTGTGCCCTCCCA
 ATACCTCTGCACACCCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCCGA
 TCTGGCCAGTGGGCACGGTCCCTACAGCTTACCCTTGGTCCCAACCCACGGTGCAACGGGA
 TTGGCGCCTCCAGACTCTCAATGGTTCCCATGCCTACCTCACCTTGGCCCTGCATTGGGTGGA
 GCCACGTGAACACATAATCCCCGTGGTGGTCAGCCACAATGCCCAGATGTGGCAGCTCCTGGT
 TCGAGTGATCGTGTGCTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGGGCCGCATGAA
 GGGCATGCCACGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAAT
 CTTCTCATCCTCATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGC
 AGACAGCGTGCCCCCTGAAGGCGACTGTCT**TGA**ATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGC
 CTCTGGCTCCATCTGAGTCCCCTGGGAGAGAGCCCAGCACCCAAGATCCAGCAGGGGACAGGA
 CAGAGTAGAAGCCCCTCCATCTGCCCTGGGGTGGAGGCACCATCACCATCACAGGCATGTCT
 GCAGAGCCTGGACACCAACTTTATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTGTG
 CCAATAATAAAGCCCCAGAGAACTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGK
ATEGPFAMDPDSGFLLVTRALDREEQA EYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHF
SQAIYRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGAL
ALSPKGSTSLDHALERTYQLLVQVKMDGDAQSGHQATATVEVSIIESTWVSLEPIHLAENLKV
LYPHHMAQVHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQA EYLLQVRAQN SHGED
YAAPLELHVLVMDENDNVPICPPRDPTVSIPELSPPGTEVTRL SAEDADAPGSPNSHVYQLL
SPEPEDGVEGRAFQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTD
INDHAPEFITSQIGPISLPEDVEPGTLVAMLTADADLEPAFRLMDFAIERGDTEGTFGLDWE
PDSGHVRLRLCKNLSYEAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQE
SYEASVPISAPAGSFLTTIQSPDPI SRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGD
TYTVLVEAQDTALT LAPVPSQYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRD
WRLQTLNGSHAYLTLALHWVEPREHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMK
GMPTKLSAVGILVGTLVAIGIFLILIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCC
TGAGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCC
TTTATCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCT
TGGTACATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAA
AGGAAATGTTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGA
GACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGGCCAGCACAGCCACCAA
CTCTGGGTCCAGTGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGAC
CTCCAATGGGGTCAGCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCAC
AGCCACCAACTCTGAGTTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTC
CAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGG
GGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGGAGGCCAGCACTGCCACCAA
CTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACT
CTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCAC
AGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTC
CAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGAACGACCTCCAATGGGGCTGGCACAGCCACCAA
CTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAGT
GTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCAC
AGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCTAGCACAGCCACCAACTCTGACTC
CAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCTCCAGTGGGGCCAAACACAGCCACCAA
CTCTGAGTCCAGTACGACCTCCAGTGGGGCCAAACACAGCCACCAACTCTGAGTCCAGCACAGT
GTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGTCAGCAC
AGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGACTC
CAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG
GATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCAAACACAGCCACCAA
CTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGAATGCACACAAC
TTCCCATAGTGCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGGTCCCTGGTGGCGTGGGA
AATCTTCCTCATCACCTGGTCTCGGTGTGGCGGCCGTGGGGCTCTTTGCTGGGCTCTTCTT
CTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCTCATGG
CCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAGGTG
GAGTCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGAA
CAGCGGGCCCCTGAAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGC
ACCAAGACCTGGTTTCCTTTTCATTTCATCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCT
GAAAATCTTGAAGAAGGTATTCCTCACCTTTCTTGCCCTTTACCAGACACTGGAAAGAGAATAC
TATATTGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCT
GTGCTTGCCCCGGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCAT
GCTGGACTCCATCTGGCATTCAAAATCTCCACAGTAAAATCCAAGACCTCAAAAAAAAAAAAA
AA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVVISSGASTATNSGSSVTSSGVSTATIS
GSSVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESS
TPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTS
SGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGA
GTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTA
TNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNS
ESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESS
TVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGG
LVPWEIFLITLVSVVAAVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAP
HRPRWSPNWFWRRPVSSIAMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCC
TACGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCA
GGTGGGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCA
ACAAATGGATGATGTGATAT**ATG**CATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCAT
GGTCAATTAACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCT
TTGGAATCATGGTGTCATGGAAAGGGATTTACTTTATACTGACTCTGTTTTGGGGAAGCTTTT
TTGGAAGCATTTTTCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATC
GCTGGATCAACAACCGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCA
TGTTTTGGTGTAAGAGTGAATTATAACTGGGGATGCATTTGTTCCCTGGAGAAAGAAGTGTCATTA
TCATGAACCATCGGACAAGAATGGACTGGATGTTCCCTGTGGAATTGCCTGATGCGATATAGCT
ACCTCAGATTGGAGAAAATTTGCCTCAAAGCGAGTCTCAAAGGTGTTCCCTGGATTTGGTTGGG
CCATGCAGGCTGCTGCCTATATCTTCATTTCATAGGAAATGGAAGGATGACAAGAGCCATTTTCG
AAGACATGATTGATTACTTTTGTGATATTCACGAACCACTTCAACTCCTCATATTCCCAGAAG
GGACTGATCTCACAGAAAACAGCAAGTCTCGAAGTAATGCATTTGCTGAAAAAAATGGACTTC
AGAAATATGAATATGTTTTACATCCAAGAACTACAGGCTTTACTTTTGTGGTAGACCGTCTAA
GAGAAGGTAAGAACCTTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACAACATTCCTC
AATCAGAGAAGCACCTCCTCCAAGGAGACTTTCCCAGGGAAATCCACTTTACGTCCACCGGT
ATCCAATAGACACCCTCCCCACATCCAAGGAGGACCTTCAACTCTGGTGCCACAAACGGTGGG
AAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAGGGGAGAAGAATTTTTATTTTACCGGAC
AGAGTGTCATTCCACCTTGCAAGTCTGAACTCAGGGTCTTGTGGTCAAATTGCTCTCTATAC
TGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCATATATTTGTACAGTCTTGTTAAGT
GGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGAGAATATTTGGTGGACTGGAGA
TCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACATTTAAATTCAAAGAAAAATG
AG**TAA**GATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTGGAAATGTTCTAAACCTT
TCTAAGCTCAGATGCATTTTTGCATGACTATGTGCAATATTTCTTACTGCCATCATTATTTGT
TAAAGATATTTTGCACCTTAATTTTGTGGGAAAAATATTGCTACAATTTTTTTTAAATCTCTGAA
TGTAATTTTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTTGGGCCAGAATA
TTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFSGIMVSWKGIYFILTLFWGSFFGSI FML
SPFLPLMFVNPSWYRWINNRLVATWLTLPVALLETMEGVKVIITGDAFVPGERSVIIMNHRTR
MDWMFLWNCLMRYSYLRLEKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYF
CDIHEPLQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLD
AVHDITVAYPHNIPQSEKHLLQGDFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERL
RSFYQGEKNFYFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITI
VIFVLQERIFGGLEIIEIACRYLLHKQPHLNSKKNE

Important features of the protein:**Signal peptide:**

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGA
 CCTCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTCATAGTGTGAGATCAACCCACAGG
 AATATCCATGGCCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTG
 GCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTG
 CTCCCTCTTTCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCCA
 TGCTGTGGTCCACCTCTACAGAGATGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCG
 AGGGAGAAGTGAAGTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTCTAAGGCTAAAAA
 CATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTACGATGAGGA
 GGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCCATCGTGGGATA
 TGTTGACGGAGGTATCCAGTTACTCTGCCTGTCTCAGGCTGGTTCCTCCAGCCACAGCCAA
 GTGGAAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAG
 CCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCAT
 CCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCA
 GCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGT
 TGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCAGGCGGAACTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCT
 GGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTGAACCCATAGAAA
 AGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCA
 GGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGACAAAATGTAGGGTGGTATGT
 GGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACAATGTGACTTTGTCTCCAACAATGG
 GTATTGGGTCCTCAGACTGACAACAGAACATTTGTATTTACATTCAATCCCCATTTTATCAG
 CCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTGGACTATGAGGGTGGGACCAT
 CTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGCTGACATGTCAGTTTGAAGG
 CTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCCATATTCAT
 ATGTCCAGTGTCTCTGGGGATTGAGACAGAGAAGACCCTGCTTAAAGGGCCCCACACCACAGACC
 CAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCTCCGGAGCCTGCGC
 ACAGAGAGTCACGCCCCCCTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCCTGAGCCCTGCA
 GCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAGTCAGAAGCC
 ATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAACCTCCATC
 CAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGGACAGT
 GATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTTTGA
 GGGCACAGTGTGTTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTTT
 GCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACAC
 CTGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTTCACTGATTCTATAAGCCCAG
 CATTACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCA
 TATCCCTCATTAACACAGACACAAAATTTCTAAATAAAATTTTAACAAATTAACTAAACAAT
 ATATTTAAAGATGATATATACTACTCAGTGTGGTTTGTCCCACAAATGCAGAGTTGGTTTAA
 TATTTAAATATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAA
 AAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAV
VHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEAT
WELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLY
DVEISIIVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVM
GMIIVFFKSKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAP
QEVPHSEKRFRTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYW
VLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLL
RPYIQHAMYDEEEKGTPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAG
GAAAAGAGTTTGTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCCT
GGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAAGAAGACCTACAATTA
CTATAGCACATTGTCATTTACAACCTGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAA
CAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAATCTCC
ATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGGAGTGTT
GGCTCATATGCTGTTGATTGTTAGATTTCACTCTACTGAGGATCCTGAAACTGTAGATAAAAT
TGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAGATCCTCA
CTCAGTTAAATTAATAAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGG
AACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGA
GGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTT
AATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAG
ATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAAT
AATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCTCTTGACAGAGCTTTC
TAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGATGCATCCTATGAGTT
TCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAATGATGGTTACAGTCA
AAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACCTCAAGC
TTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGC
ATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGC
TGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTTTATACTAGAGT
TACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAA
CAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATTTTGTAGAGATACAGAATTGGAGA
AGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAACTGTTTGCTTGATGCATGTATTTT
CTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCAGATCAACTCTGTCATCTGTGA
GCAATAGTTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTACATTACAGCCTGTA
TTCATTTGTTCTCTAGAAGTTTTGTGAGAATTTTGAAGTTGTTGACATAAATTTGTAATGCATA
TATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTTCAGCAAATATCCA
TTTTCAAGGTGCAGAACAAAGGAGTGAAAGAAAATATAAGAAGAAAAAAATCCCCTACATTTTA
TTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTCATTAT
GAAAGGTCAAGCAAAGACAGCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACTAA
GTTAAGGAAGTCCAGAAAGAAGCCAAGATATATCCTTATTTTCATTTCCAAACAACCTACTATG
ATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCTATAATAATTATACAACTTCATGCAATG
TACTTGTTCTAAGCAAATTAAGCAAATATTTATTTAACATTGTTACTGAGGATGTCAACATA
TAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDK
LYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFH
STEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQS
LRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIK
PSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTG
FGALKNDGYSONHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLV
SSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCC
TGCCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACC GGCC
CAGCGCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTC
TCCCTTACGGGGCTCACA**ATG**GCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTAC
GCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGG
GACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTG
ACTTACTTTTCTGTGGTTCATCCGGTCATGATTGCTGTTTGCTGTTTCCCTTATCATTGTGGGG
ATGTTAGGATATTGTGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGT
TTGCTTGTCATTTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATG
GTTCCAGTACAATGGTCAGATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCT
AGATATCGGTGGCTTACTCATGCTTGGAAATTTTTTTTCAGAGAGAGTTTAAGTGTGTGGAGTA
GTATATTTCACTGACTGGTTGGAAATGACAGAGATGGACTGGCCCCCAGATTCCCTGCTGTGTT
AGAGAATTCCCAGGATGTTCCAAACAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAG
GGTTGTGGGAAGAAAATGTATTCCTTTTTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTT
CTGGGAATCTCCATTGGGGTGACACAAATCCTGGCCATGATTCTCACCATTACTCTGCTCTGG
GCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAATGATGCTCCTTGAAGAATGACAAC
TCTCAGCACCTGTCTGTCCTCAGTAGAACTGTTGAAACCAAGCCTGTCAAGAATCTTTGAA
CACACATCCATGGCAAACAGCTTTAATACACACTTTGAGATGGAGGAGTTAT**TAA**AAAGAAATG
TCACAGAAGAAAACCACAAACTTGTTTTTATTGGACTTGTGAATTTTTGAGTACATACTATTGTG
TTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAATAACACCTAAGCATATACTATTCTA
TGCTTTAAATGAGGATGGAAAAGTTTCATGTCTATAAGTCACCACCTGGACAATAATTGATGC
CCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCCTGTGTATGACTTTTACTGAAC
ACAGTTATGTTTTGAGGCAGCATGGTTTTGATTAGCATTTCGCATCCATGCAAACGAGTCACA
TATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAACTAGTATATAAAGTACT
AATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTACTCAGCGATCTATTC
TTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACTACCTAAATGTGAT
TTTTGCTGGTTACTAAATATTCTTACCCTTAAAAGAGCAAGCTAACACATTGTCTTAAGCT
GATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAAGTCGATTTCAAGTTCT
GATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATTATTTTTTA
GCCTTTCCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAACTGTTA
TTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTCAGA
ATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTCAGAA
AGGACTTGATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTTAAAA
AGCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTCTACAGA
AAATAGTGTTCTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAACATGTGACAATTTAG
AGATTCTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTT
TTTACAAGAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTAT
TTTGTATTATTTCTCAGAATATGGAAAGAAAATTAAAATGTGTCAATAAATATTTTCTAGAGAG
TAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVV
HPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWS
DMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGC
SKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDR
REPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

108/168

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCATCCTCTTCCTCTGATAAAGCCC
CTACCAGTGCTGATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAAAGTGCTTGA
AAGAGAAGGGGACAAAGGAACACCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTC
CAGAAGG**ATG**CCTCCATTCTGCTTCTCACCTGCCTCTTCATCACAGGCACCTCCGTGTCACC
CGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCA
CCAGTTGGATGAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCA
CTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAAC
CCACGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGGCGACGGCATTGTGCAACGCCA
GGCTTGTGCCAGCTTCAATGGGAACCTGCTGTCTCTGGAACACCACGGTGGAAAGTCAAGGCTTG
CCCTGGAGGCTACTATGTGTATCGTCTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGG
TCATTTTTATGACATCTGCGACGAGGACTGCCATGGCAGCTGCTCAGATACCAGCGAGTGCAC
ATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGCTTTGATGAAAATGAATGTGA
GCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCCTACCGCTGTGAGTG
TGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAA
TAACAATGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCC
CCGGGGCCTGGTGCTGTCTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAATCAAA
TGCCATTGAAGTGAACATCCCCAGGGAGCTGGTTGGTGGCCTGGAGCTCTTCTGACCAACAC
CTCCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACATCCTCTTCTCTCAAGACATGTGG
TACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGACAGGTCTACCCAA
GCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTGAC
CTGCGAGTTTCCACGCCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCC
ACTGGAAATCATGAGCCGAAATCATGGGATCTTCCCATTCCTCTGGAGATCTTCAAGGACAA
TGAGTTTGAAGAGCCTTACCGGGAAGCTCTGCCCACCCTCAAGCTTCGTGACTCCCTCTACTT
TGGCATTGAGCCCGTGGTGCACGTGAGCGGCTTGGAAGCTTGGTGGAGAGCTGCTTTGCCAC
CCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGGATGGCTGTGTTTCAGA
TGAATCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGGTCCCTGTCTT
CAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGT
GTTGGACGAGCGTTCCTCGCTGTGCCCAGGGTTGCCACCGGCGAATGCGTCGTGGGGCAGGAGG
AGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCGACTGGGA
GGACT**TAG**TTTCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTC
CCCCACCGCCCTCTAAGAACATCTGCCAACAGCTGGGTTTCAGACTTCACACTGTGAGTTTCAG
ACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTTCAGTGGGCACAGGTACACAGCACTGCTG
AACAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAACTGTCCACCCAGAA
AGACACTACCCCATTTCCCTCATTTCTTTCTTACACTTAAATACCTCGTGTATGGTGAATC
AGACCACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAG
TTACTGAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTT
GAAATTTCAATTCAAATGCAGACTAATTATAGGGAATTTGGAAGTGATCAATAAAACAGTAT
ATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFT
 GMAGDAMPTFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPG
 GYYVYRLTKPSVCFHVYCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQN
 NGGCSEICVNLKNSYRCECGVGRVLRSDGKTCEDEVEGCHNNNGGCSHSCLGSEKGYQCECPRG
 LVLSEDNHTCQVPVLCKSNAIEVNI PRELVGGLELFLTNTSCRGVSNNGTHVNI LFSLKTCGTV
 VDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLE
 IMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKL RDSLYFGIEPVVHVSGLES LVESCFATPT
 SKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRVLCGVLD
 ERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGP I RIDWED

Important features of the protein:**Signal peptide:**

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
 522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCT
 GCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACC
 TGTGTGGGGAGGCCCTCCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACC
 GGGAGGATCACAGAGCCAGCATGTTACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCG
 ATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCA
 TCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGTGGTTGTCTCATCAAGGTGATTC
 TGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTG
 ACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAG
 GGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCA
 CAGGGAAGTGGTTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTA
 GGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAA
 TCACAGAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCC
 TGGTCTCCCTGCACTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGG
 AGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAGTACGACAAACAGCACGTCT
 GTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAACATA
 CCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAGCTTCCCATCCCTGG
 CTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGACATCGCCC
 TCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCT
 TTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGA
 ATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGT
 GCAATGCAGACGATGCGTACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGG
 AAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGCCCTGATGTACCAATCTGACCAGT
 GGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGCCCGAGCACCCCAGGAGTAT
 ACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTGTAATGCT
 GCTGCCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCTGGGGATCCCCCA
 AGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCATTTCTT
 GGAGCAGCAAAGGGCCTCAATTCTGTAAAGAGACCCTCGCAGCCCAGAGGCGCCCAGAGGAAG
 TCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACT
 GAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGC
 AGGCTGTCTTGTAAGGCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCA
 GCCCTGTCCGTCTTCACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAAAAT
 GCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATGG
 CCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAAAAAAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFL
CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSA
CFDNFTEALAETACRQMGYSRAVEIGPDQDLVDVEITENSQELMRNSSGPCLSGSLVSLHCL
ACGKSLKTPRVVGGEESVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWK
VRAGSDKLGSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTP
ATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTC
QGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGAT
TATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTTCAGCAACTAAAAAGCCACAG
GAGTTGAACTGCTAGGATTCTGACT**ATG**CTGTGGTGGCTAGTGCTCCTACTCCTACCTACATT
AAAATCTGTTTTTTGTTCTCTTGTAAGTAGCCTTTACCTTCCTAACACAGAGGATCTGTCACT
GTGGCTCTGGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGT
CCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAAC
TGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGG
CTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGAT**TGA**CAAGCCCGAAGATTTTCATAGGCG
ATGGCTCCCACTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGG
GCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTT
CATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCT
TGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGA
CTTGCAATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTG
TGCCAAATTATGGGTGAGAAAAGATGGAGGTGTTGGGTATCACAAGGCATCGAGTCTCCTGC
ATTCAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACTCGGGACTCACCTCTG
GGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTGAAGGGCAACTGCAGGC
CGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACCAGCCAGGGGCAGCCGTCTG
GGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCTCCTTCCCTCTGAGAGGCCCTCCTATGT
CCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGCTAATGGCTCAGTGTGGCCCAGGA
GGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCGAACACGGAAATGCCTCCAGT
AAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTTAAATCATGTTCTAGT
AATTGGAGCTGTCCCAAGACCAAAGGAGCTAGAGCTTGGTTCAAATGATCTCCAAGGGCCCT
TATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCAAACCTAAGAACCAGGTGCA
TTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACATTTTGGGAGGCCGAGGCG
GGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCTGTCTC
TACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATCCCAGCTACTCGGG
AGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACAGGAGAATCACT
TCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTATGGTTATTT
GTAA

114/168

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTAS
PCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

114/168

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACC**ATGGC**
AAAGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAA
GAAAATATGTAAATCACTTAAGATTTGTGGACTGGTGTTTGGTATCCTGGCCCTAACTCTAAT
TGTCCTGTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGA
GCACACTTTCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAG
AACTGAAATATTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAA
CGGATACACTGGCATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGT
GATTCCTGAATTTTCTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTT
CTTTGAACAGTCAGTGATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAA
AAATTCCAAAATTCTGGAGATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAAT
ATCAGTTTCTGAGTTACAAGACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGA
AAAAAAGGGATTGAACAAAATGAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCG
TCACGCCAGACAAGCAAGTGAGGAAGAACTTCCAATAAATGACTATACTGAAAATGGAATAGA
ATTTGATCCCATGCTGGATGAGAGAGGTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTA
TTGCCGCCGCGTCTGTGAACCTTTACTAGGCTACTACCCATATCCATACTGCTACCAAGGAGG
ACGAGTCATCTGTCGTGTCATCATGCCTTGTAAGTGGTGGGTGGCCCGCATGCTGGGGAGGGT
CTAATAGGAGGTTTGAGCTCAAATGCTTAAACTGCTGGCAACATATAATAAATGCATGCTATT
CAATGAATTTCTGCCTATGAGGCATCTGGCCCTGGTAGCCAGCTCTCCAGAATTACTTGTAG
GTAATTCCTCTCTTCATGTTCTAATAAACTTCTACATTATCACCAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTILIVLFWGSKHFWPEVPPKAYD
 MEHTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIIYFVGLQKCFIKTQI
 KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPT
 LISVSELQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENG
 IEFDPMLDERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWWVARMLGRV

Important features of the protein:**Signal peptide:**

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGC
AGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCCCTCC
TGTCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGG
ACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGA
GGCAGAGTTCAGGCTTCACCGAATGCAGGCCCTATTTACCATCCTGGGACTTCCAGCCATGC
TGCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTAT
CCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACATGA
CACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGT
TTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTG
GGATGGTGCAGACTGTTCCAGACCAGGTACACATTTGGTGCGGCTCTGTTTCGTGGGCTGGGTG
CTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAG
AAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTCCTACAAGCCTG
GAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAGATATACGATGGAG
GTGCCCCGCACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTATGTG**TAAT**GCTCTA
AGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGATCCCA
TCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTG
GAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAG
TTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTTAAATATAA
CTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAG
ACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAG
AAAACTTTTGAAAGGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTGAATTTTGTCTC
CCCACCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATT
TGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAGTTACCAAACCA
AAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTGTTGACATCTTCTTATTACAGC
AACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCCTCTTCTGTCGCGGGTCAGAAA
TTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAATATAGTTAAAATAA
ATAATGTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAG
AAGGAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAA
GTACAAATTCCATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGG
ATCACTTGAGCCCAGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAA
ATACAGAGAGAAAAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTCGGGGAG
GCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGTGGGGCTGCAGTGAGCCATGATCACACC
ACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGA
ACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLRSCVRQSSGFTEC
RPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFI
VSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGG
VMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQ
SYPSKHDYV

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGG
 AGTCCAGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGG
 CTGTTTCTTGGTGGTGGTGGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGA
 GTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAGAACTTCTGGGAAGGACTGTGGATG
 AATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTT
 TCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCT
 TTCTATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAAGGTGAAG
 GCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGTCTATCCCT
 GTGAGCTGGGTGGCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAA
 AAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGTCTGATTGTTGGA
 GGAGCTCTGTTCTGCTGCGTTTTTTGTTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATA
 CCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCGTCTACTCC
 AGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATGACA
 AAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAACTGC
 CTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGCAGAA
 TGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGG
 TTCAAGCATCTACTCTTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTT
 ACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCA
 CATAGAGACATGCTTATATGGTTTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAAT
 AGAACTCAACTATTGCTTTTTCAGGGAAATCATGGATAGGGTTGAAGAAGGTACTATTAATTG
 TTTAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGAAGATTAAATGAAGGCTTTAA
 TCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTGTTTTTTAGCCTAGGAGTTAGAA
 ATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTTCTTGTGTATTAAATTAACATT
 TTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAAGTCTTTTCCAGGGCTATACTC
 AGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTGAAAATATTTTT
 GTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGATATATTT
 TAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATAT
 GTCTTGGTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTC
 ACCTGCTCCTATGTGGGTACCTGAGTCAAATTTGTCATTTTTGTTCTGTGAAAAATAAATTC
 CTTCTGTACCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTTATT
 CCAAATTTGATGAACTGACAATCCAATTTGAAAGTTTGTGTGCGACGTCTGTCTAGCTTAAAT
 GAATGTGTTCTATTTGCTTTATACATTTATATTAATAAATTTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQ
CKIYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIF
IITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCN
EKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

120/168

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGG
AGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGC
AGCCGGGAGCCATGCGACCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGC
TGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAA
AGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAG
CAGGAGTGCCTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCGGGTACACCTGGGATCC
CAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCT
GGACACCCAACTACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAA
TTGCGGAGTGTACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCT
CACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTG
AATGTTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGA
ATTCAACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTG
GATTAGTGGATGTTGCTATCTGGGTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTA
CTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAATAAATGCTTTAATTTT
CATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAATGACATTTTAAATAAG
TTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGACCAAAGTGTGATTTC
CACTGTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAAGTGGTTTCAATATTTTTTT
TAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTATAATTTGGAATATTGTTG
TGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTTAAAAAATATAAAAGCTACCAATCTTTG
TACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGTAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAAGVP
GRDGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAEC
TFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTI
NIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

122/168

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTG
AACTGGGTGCTCATCACGGGAAGTGTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCC
CCAAATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTT
TTTTTAACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCAT
GAAGATCCTATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTT
GTTCTTGGAGTGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAG
GGGTCCAATTTTTCTTCCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGG
CTGTCATGCAACTGGCCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAA
AGGATGGGTTTCAATGTAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCC
ACTGTCTTACTGACAATGCTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAA
GGCAAAATGGTATATTGTGAATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGT
TGCTTAGGTTTGTCCCTTCGCTATAACAGCCTTCAAAAACCTAAGTATAATCAATTTAAAGGG
CTCAACCAGCTCACCTGGCTATACCTTGACCATAACCATATCAGCAATATTGACGAAAATGCT
TTTAATGGAATACGCAGACTCAAAGAGCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTT
AACAAATACCTTCAGACCTGTGACAAATTTACGGAACCTTGGATCTGTCCTATAATCAGCTGCAT
TCTCTGGGATCTGAACAGTTTCGGGGCTTGCGGAAGCTGCTGAGTTTACATTTACGGTCTAAC
TCCCTGAGAACCATCCCTGTGCGAATATTCCAAGACTGCCGCAACCTGGAACCTTTTGACCTG
GGATATAACCGGATCCGAAGTTTAGCCAGGAATGTCTTTGCTGGCATGATCAGACTCAAAGAA
CTTCACCTGGAGCACAATCAATTTTCCAAGCTCAACCTGGCCCTTTTCCAAGGTTGGTCAGC
CTTCAGAACCTTTACTTGCAAGTGAATAAAAATCAGTGTGATAGGACAGACCATGTCCTGGACC
TGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGATCGAAGCTTTCAGTGGACCCAGT
GTTTTCCAGTGTGTCCGAATCTGCAGCGCCTCAACCTGGATTCCAACAAGCTCACATTTATT
GGTCAAGAGATTTTGGATTCTTGGATATCCCTCAATGACATCAGTCTTGCTGGGAATATATGG
GAATGCAGCAGAAATATTTGCTCCCTTGTAAGCTGGCTGAAAAGTTTTAAAGGTCTAAGGGAG
AATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTGATCGATGCAGTGAAG
AACTACAGCATCTGTGGCAAAGTACTACAGAGAGGTTTGATCTGGCCAGGGCTCTCCCAAAG
CCGACGTTTAAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTTTGCCCCGACG
GTGGGAGCCACAGAGCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCTTTCCATAAA
ATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCCTGCTGGTTATCTACGTG
TCATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAAGGCAC
AGGAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTTTATGTAGAT
TATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTAT
AACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTTGTGATAAAAAGAGCTCTTAAAAGC
TGGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTC
CCCTTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCCTTGTCGTTTTAGTGCAATTCATA
ATACTGGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGT
GAAGCTTGAACCTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAAT
GTCGCATTTGTTTTAAGATAAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMVYCESQKLQEIPSSISAGC
 LGLSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLN
 NTFRPVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLG
 YNRIRSLARNVFAGMIRLKEHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSTW
 SSLQRLDLGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWE
 CSRNICSLVNWLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKP
 TFKPKLPRPKHESKPPPLPTVGATEPGPETDADAHEHISFHKIIAGSVALFLSVLVILLVIYVS
 WKRYPASMKQLQQRSLMRRHRKKKRQSLKQMTPTSTQEFYVDYKPTNTTETSEMLLNGTGPCTYN
 KSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGA**ATG**TCCGTCCGGAGGAGGAGGAGAGGCTTTTGCCGC
TGACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCG
AGCTAGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTC
TTGCTCGGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCC
TAGGGATCATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACA
GACACGTAGTGTATTCTGGAGGTGCAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTG
GCAAAAGTGAAGATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTG
TTATTGGCCAGTTTTTAGCCAATCCAACCTGACCTAGTGAAGGTTTCAGATGCAAATGGAAGGAA
AAAGGAACTGGAAGGAAAACCATTTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCT
TAGCTGAAGGAGGAATACGAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCAC
TGGTGAATATGGGAGATTTAACCCTTATGATACAGTGAAACACTACTTGGTATTGAATACAC
CACTTGAGGACAATATCATGACTCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTA
TTCTGGGAACACCAGCCGATGTCATCAAAGCAGAATAATGAATCAACCACGAGATAAACAAG
GAAGGGGACTTTTGTATAAATCATCGACTGACTGCTTGATTTCAGGCTGTTCAAGGTGAAGGAT
TCATGAGTCTATATAAAGGCTTTTTACCATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGT
TCTGGCTTACTTATGAAAAATCAGAGAGATGAGTGGAGTCAGTCCATTT**TAA**

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGAR
ESAPYRGMVRTALGIIIEEGFLKLWQGVTPIYRHVVYSGGRMVITYEHLREVVFVKSEDEHYP
LWKSVIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFKILAEGGIRGL
WAGWVPNIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVI
KSRIMNQPRDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIR
EMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

126/168

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGCAGGAGAGCGGCCGGGCGTCTAGCTCCTCG
ACCCCCGTGTCTGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCC**ATG**GCCAGGCCCGGC
ATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATCGGCGCGGGCC
GTGGCCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAAC
ATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGA
TGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGC
GGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCCGGCCTGACACCCTGCTCTCAGGCAGC
ACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAA
GCCTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATG
TCTGGCCACCGAGTGTTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTC
ACTGCGCTGACAGAGGGACTGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACG
TGCATCTCTCCAGGTGTGGTGGAGACACAATTCGCCTTCAAACCTCCACGACAAGGACCCTGAG
AAGGCAGCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTT
ATCTACGTCCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCCCACGGAG
CAGGTGACC**TAG**TGACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTG
CCTCTGGATTTTAGGTGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCGACC
AGGGGCTAGAAAATTTGTTTGAGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATG
TGAAAAATGGGCTGGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTAACTTGTTCCTG
TGCCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTT
GTGGCCAAAATCCCCATCTTCTTGCACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG
GAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCC
ACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCAGTCTTGGCTTCTTGT
CCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC
CCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAAATCGCAACCAAAAAAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGT
LIPYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALS
ICTREAYQSMKERNVDDGHIININSMGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQT
HIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQ
MRPTEQVT

Important features of the protein:**Signal peptide:**

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACC
ATCATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGG
GAGCACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTT
TCATGCCAGCCTGTAAAAGGCCATGGAACCTTGGGTGAATCACCGATGCCATTTAAGAGGGTT
TTCTGCCAGGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAG
CCACCTGTGGCCGTTGAGTGCTTGAAATTGAGGAACTGAGAAAATTAATTTCTCATGTATTTTT
CTCATTTATTTATTAATTTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGG
ATACATGTATACAATATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACAT
TTATTTTTTATTCTTTTTTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCC
ATCTCAGCTTACTGCAACCTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAA
GTAGCTGGGACTACAGGCATGCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACG
GGGTTTTGCCATGTTGCCCAGGCTGGCCTTGAACCTCCTGGCCTCAAACAATCCACTTGCCTCG
GCCTCCCAAAGTGTTATGATTACAGGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTT
CTTTGTGTTGGGAACCTTTGAAATTATACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTG
CTATGGAACACTGGGACTTCTTCCTCTATCTAACTGTATATTTGTACCAGTTAACCAACCGT
ACTTCATCCCCACTCCTCTCTATCCTTCCCAACCTCTGATCACCTCATTCTACTCTCTACCTC
CATGAGATCCACTTTTTTAGCTCCCACATGTGAGTAAGAAAATGCAATATTTGTCTTTCTGTG
CCTGGCTTATTTCACTTAACATAATGACTTCCTGTTCCATCCATGTTGCTGCAAATGACAGGA
TTTCGTTCTTAATTTCAATTAAAATAACCACACATGGCAAAAA

130/168

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRI PVLVSCQ
PVKGHGTLGESPMPEKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAG
 CCTGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTCAAGTAGCACCTCTATCAG
 TTATGGCTAAATCCTGTCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATC
 GCTTTCTGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAACA
 ACCAAATAAATAATGCTGGGATTTCCTTCAGATTTGAAAACTTGCTGAAAGTAGAAAGAATAT
 ACCTATACCACAACAGTTTAGATGAATTTCTACCAACCTCCCAAAGTATGTAAAAGAGTTAC
 ATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAAAAATTCCCTATCTGG
 AAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCATTCGAG
 ACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTT
 TGCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCAT
 CTCTTCAAGGTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATG
 GTTTAGGTGACAAAGTTTTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATT
 CCCTGACTGCTGCACCAGTAAACCTTCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATA
 ACCACATCAATCGGGTGCCCCCAATGCTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATA
 TGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTTGATGATTTGGACAATATAACAC
 AACTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGT
 TACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAAAGGTTT
 GTGGGATGGCTATTAAGGATCTCAATGCAGAAGTGTGTTGATTGTAAGGACAGTGGGATTGTAA
 GCACCATTCAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCAAGGACAGTGGCCAG
 CTCCAGTGACCAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAG
 GGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCTATA
 TCTCTTGAAACTTGCTCTACCTATGACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATA
 GCCCGGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAACGCAGTGAGTACTTGGTCA
 CAGCCCTGGAGCCTGATTCACCCTATAAAGTATGCATGGTTCCTATGGAAACCAGCAACCTCT
 ACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTTCGAATGTACAACC
 CTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACCTTTGG
 CTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGT
 ATGTTTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAA
 AGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACTT
 CTTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAATACACACCA
 TATTTCTCCTAATGGAATGAATCTGTACAAAACAATCACAGTGAAAGCAGTAGTAACCGAA
 GCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACTCATGATGCTGAAGGACTCACA
 GCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATGGT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTL
 YLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNI RTITYDSLK
 IPYLEELHLDDNSVSAVSIEEGA FRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRIST
 ISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKL
 YLQDNHINRVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDL DNITQLILRNNPWYCGCKMKW
 VRDWLQSLPVKVNVRGLMCQAPEKVRGMAIKDLNAELFDC KD SGIVSTIQITTAIPNTVYPAQ
 GQWPAPVTKQPDIKNPKLTKDQQTGSPSRKTITITVKS VTSDTIHISWKLALPMTALRLSWL
 KLGHSPAFGSITETIVTGERSEYLVTALEPDSPYKVCMPMETS NLYLFEDETPVCIETETAPL
 RMYNP TTTLNREQEKEPYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSK
 GRRRKDDYAEAGTKKDNSILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNL YKNNHSES
 SSNRSYRDSGIPDSDHSHS

Important features of the protein:**Signal peptide:**

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
 640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCCAGGCTTCTTGG
 CAGCCCTGCCGGGCCACTTGTCTTC**ATG**TCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAG
 GGCGTGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGA
 GCAGAGGCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACC
 GCTGTGGTCCCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCT
 GATCCTGTCTGCGCTGGAGAGAGCCACCGTCTTCCCTAGAACAGAGGCTGCCTGAAATCAACCT
 GGATGGCATGGTGGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGC
 CCAGGAGCCCCCTGCTGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGC
 TGCCATCCAGAGATCCCTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCA
 GCTGACCCTCCAGCCCGGGTTTTTGGAAAGCTCCCACATGCCTGGATCCACACTGATGCCTCCTT
 GGTGTACCCACGTTCCGGGCCCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGT
 GCAGCTGCTGGGAACCGGGACGGACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAG
 CCTCATGACCAAGCCCGGCTGCTCAGGCTACTGCCTGTCCCACCAACTGCTCTTCTTCCTCTG
 GGCCAGAATGAGGGGATGCACACAGGGACCACTCCAACAGAGCCAGGACTATATCAACCTCTT
 CTGCGCCAACATGATGGACTTGAACCGCAGAGCTGAGGCCATCGGATACGCCTACCCTACCCG
 GGACATCTTCATGGAAAACATCATGTTCTGTGGAATGGGCGGCTTCTCCGACTTCTACAAGCT
 CCGGTGGCTGGAGGCCATTCTCAGCTGGCAGAAACAGCAGGAAGGATGCTTCGGGGAGCCTGA
 TGCTGAAGATGAAGAATTATCTAAAGCTATTCAATATCAGCAGCATTTTTTCGAGGAGAGTGAA
 GAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGTTGCTCAGGCTGGAGTACAGTGGCGCAA
 TCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGCAATTCTCTTGCCTCATCTCCCGAG
 TAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAATTTTTTATATTTTTTTAGTAGAGAC
 AGGGTTTCATCATGTGCTCATGCTGGTCTCGAACTCCTGATCTCAAGAGATCCGCCCACCTC
 AGGCTCCCAAAGTGTGGGATTAT**TAG**GTGTGAGCCACCGTGTCTGGCTGAAAAGCACTTTCAAA
 GAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTCATGGGGGCTCTCTCCCC
 TAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGGGTGGCTTCTCTATACA
 TCCTGGCAGAATACCCCCCAGCAAACAGAGAGCCACACCCATCCACACCGCCACCACCAAGCA
 GCCGCTGAGACGGACGGTTCATGCCAGCTGCCTGGAGGAGGAACAGACCCCTTTAGTCCTCA
 TCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGGATAAGCAAA
 GCCACCCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGGGGCCGGG
 AGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLLARAQGAERVTSSEQRPAMASLGLLLLLLLTALPPLWSSSLPG
LDTAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPL
SLRVGMLGEKLEAAIQRSYLHKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQ
DSFSEERSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQ
GPLQQSQDYINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILS
WQKQQEGCFGEPPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLP
PGFKQFSCILIPSSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:**Signal peptide:**

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTC**ATG**GCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGC
TGCTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGG
GAGGGAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGG
TGAAGCCTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTT
TCCTTAAGACAGATGGGAGTTTTGTGGTTTCATGATATACCTTCTGGATCTTATGTAGTGGAAG
TTGTATCTCCAGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGA
GAGCAAGATATGTGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAA
TGAAATCTTCAGGTCCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTC
TAATGAACCCAATGGTTATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAG
TGGTCAACACAAGTGATCCTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATT
CCAACCATGAGTTGCCTGATGTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTG
GCAAATCTAGCAGCGGCAGCAGTAAAACAGGCAAAAGTGGGGCTGGCAAAGGAGG**TAG**TCAG
GCCGTCCAGAGCTGGCATTTCACAAAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAAAA
CCGTGTGAAGCAACTACTATAAACTTGAGTCATCCCGACGTTGATCTCTTACAACACTGTGTATGTT
AACTTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGT
ATGAGGTCAATATTGATGTCAGTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAA
TTATATGAACTACTATACATTATGTATATTAATTAAAACATCTTAATCCAGAAATCAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVVPGVKPDWIS
AARVLVDGEEHVGFLLKTDGSEFVVHDIPSGSYVVEVVSAPYRFDVPRVDITSKGKMRARYVNYI
KTSEVVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPD
MRREMEQSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:**Signal sequence:**

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCTGTGGGTCTGAGGG
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAA
ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTACAGGCCTAAGATGAAAGCCTCTAGT
CTTGCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACA
TGCCATTGTGGGGAGGAAGCAATGAAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTG
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAACTAGACATTCTTCTGCAATGGATGGAG
GAGACAGAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTCGAGGTCAAGAGCTCCAGTCT
TCAATACCTGCAGAGGAGGCATGACCCCAAACCACCATCTCTTTACTGTACTAGTCTTGTGCT
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT
TATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTACTTGGACATG
AACTTTAAAAAAATTCACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT
ACAGTAAAAAAAACCTTGTAATTCTAGAAGAGTGGCTAGGGGGGTATTTCATTTGTAT
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGAATTGCACATCTACCTTACAATTACTG
ACCATCCCCAGTAGACTCCCCAGTCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG
CCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPE
IFSSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQE
IRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHY
TLRKISSLANSELTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDI
LLQWMEETE

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTCGGTTCGCGATCCAG
CCTAGCGTGTCCACG**ATG**CGGCTGGGCTCCGGGACTTTTCGCTACCTGTTGCGTAGCGATCGAG
GTGCTAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCCTCTGCCAGA
GCGGAACACGGAGCGGAGCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTGGACC
ACGCTGCCACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGAT
TTTGTGTTTGGGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGA
GCATCTCACAGTTTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCA
TTGATGACGGGGAGCCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTG
CTGGAAGACAGTGTGATAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGAT
GAAACCTGGGTAAATTATTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTTC
GTGTCAGATTACACAGAGGTGGATAATAATGTCACGAGGCATTTGGATAAAGTATTA AAAAGA
GGAGATTGGGACATATTAATCCTCCACTACCTGGGGCTGGACCACATTGGCCACATTTTCAGGG
CCCAACAGCCCCCTGATTGGGCAGAAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCAC
ACCTCACTGCAGTCGAAGGAGAGAGAGACGCCCTTACCCAATTTGCTGGTTCTTTGTGGTGAC
CATGGCATGTCTGAAACAGGAAGTCACGGGGCCTCCTCCACCGAGGAGGTGAATACACCTCTG
ATTTTAATCAGTTCTGCGTTTGAAAGGAAACCCGGTGATATCCGACATCCAAAGCACGTCCAA
TAGACGGATGTGGCTGCGACACTGGCGATAGCACTTGGCTTACCGATTCCAAAAGACAGTGTA
GGGAGCCTCCTATTCCAGTTGTGGAAGGAAGACCAATGAGAGAGCAGTTGAGATTTTTACAT
TTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAGAATGTGCCGTCATATGAAAAAGATCCT
GGGTTTGAGCAGTTTAAAATGTCAGAAAGATTGCATGGGAAGTGGATCAGACTGTACTTGGAG
GAAAAGCATTGAGAAGTCCTATTCAACCTGGGCTCCAAGGTTCTCAGGCAGTACCTGGATGCT
CTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAGTTCTCACCTGCTCCTGCTCA
GCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCACTGTCATCTCCTGGGTTTT
CTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTCGGCCGTTACGTCATTGTGTGCACCTCAG
CTGAAAGTTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCTGCCTTTCGTTTACCA
GACTCTGGTTGAACACCTGGTGTGTGCCAAGTGCTGGCAGTGCCCTGGACAGGGGGCCTCAGG
GAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCGACACAGGTGTTACATCTGT
GCTGTCAGGTCAGATGCCTCAGTTCTTGGAAGCTAGGTTCTGCGACTGTTACCAAGGTGAT
TGTAAGAGCTGGCGGTCACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAATCGGACA
GCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGACACTC
AGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCTGCC
TCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTTGGAATTCTTAGTCCTTGGCCT
CGGACACCTTCATTTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCA
CACTCAGATCCACAGAGCCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCC
CCCACCCCAACCCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTCAGAGGCCCTGTGCTG
AGTGTCTGACCGAGACACTCACAGCTTTGTCATCAGGGCACAGGCTTCCTCGGAGCCAGGATG
ATCTGTGCCACGCTTGACCTCGGGCCCCTCTGGGCTCATGCTCTCTCCTGCTATTGAATT
AGTACCTAGCTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPP
LFSKVIVLIDALRDDDFVFGSKGVKFMPTTYLVEKGASHSEFVAEAKPPTVTMPRIKALMTGS
LPGFVDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFEYDGTTSFFVSDYT
EVDNNVTRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQS
KERETPLPNLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

GGCACGAGGCAAGCCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTA
CAGAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTG**ATG**TTACTGCTGCTGTTGGAGTAC
AACTTCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAAC
TTAAACCCGAAGAAATTCAGCATTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGG
AATCTCATAGCAGTTCCAGATAAAAACCTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCA
TCCTTGAGCTCAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAG
TTTTGTCTCTACTGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAG
AAACTGATGAAGCTGGCTGCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCT
CAGGTGGGCTCCTGGAACATGCTGGAGTCGGCGGCTCACCCGGATGGTTCATCTGCACCTCC
TGCAATTGTAATGAGCCTGTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTT
TCATTTCAACCAGTTTGCAAAGCTGAAATGAGCCCCAGTGAGGTCAGCGAT**TAG**GAAACTGCC
CCATTGAACGCCTTCCTCGCTAATTTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVPDKNYIRP
EIFFALASSLSSASAEKGS PILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESAR
RPFIFYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPS
EVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

Year	Month	Day	Time	Location	Observer	Remarks
1961	July	1	10:00	1000 ft	W. H. H.	1st
1961	July	1	10:00	1000 ft	W. H. H.	2nd
1961	July	1	10:00	1000 ft	W. H. H.	3rd
1961	July	1	10:00	1000 ft	W. H. H.	4th
1961	July	1	10:00	1000 ft	W. H. H.	5th
1961	July	1	10:00	1000 ft	W. H. H.	6th
1961	July	1	10:00	1000 ft	W. H. H.	7th
1961	July	1	10:00	1000 ft	W. H. H.	8th
1961	July	1	10:00	1000 ft	W. H. H.	9th
1961	July	1	10:00	1000 ft	W. H. H.	10th
1961	July	1	10:00	1000 ft	W. H. H.	11th
1961	July	1	10:00	1000 ft	W. H. H.	12th
1961	July	1	10:00	1000 ft	W. H. H.	13th
1961	July	1	10:00	1000 ft	W. H. H.	14th
1961	July	1	10:00	1000 ft	W. H. H.	15th
1961	July	1	10:00	1000 ft	W. H. H.	16th
1961	July	1	10:00	1000 ft	W. H. H.	17th
1961	July	1	10:00	1000 ft	W. H. H.	18th
1961	July	1	10:00	1000 ft	W. H. H.	19th
1961	July	1	10:00	1000 ft	W. H. H.	20th
1961	July	1	10:00	1000 ft	W. H. H.	21st
1961	July	1	10:00	1000 ft	W. H. H.	22nd
1961	July	1	10:00	1000 ft	W. H. H.	23rd
1961	July	1	10:00	1000 ft	W. H. H.	24th
1961	July	1	10:00	1000 ft	W. H. H.	25th
1961	July	1	10:00	1000 ft	W. H. H.	26th
1961	July	1	10:00	1000 ft	W. H. H.	27th
1961	July	1	10:00	1000 ft	W. H. H.	28th
1961	July	1	10:00	1000 ft	W. H. H.	29th
1961	July	1	10:00	1000 ft	W. H. H.	30th
1961	July	1	10:00	1000 ft	W. H. H.	31st

AAAAAAAAAAAAAAAAAAAA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSK
DYYAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTI
STRPWMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCG**ATG**AGGA
AGCACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGG
TCCAGACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCCAGCACTGCCC
AGATCACTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGCAAGC
TCGACATTGACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCG
ATGGCATCCACTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCT
GCATCAATGCCACCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACC
AGCAGGTGCTCTGGCGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGG
AGAGGGGCGCAGGACTTCGGGTCAACATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGA
TCTGGCTCATGGTGAAAT**TAA**GCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAA
TCCTGGCAAGTGACCCAGCTCTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGC
GGCGATGCACTCGCACTGCAAATGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCT
GATAGATGGGGGACTGTGGCTTCTCCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCA
CACTAGATTAGTAGTAAATGCTTGATGAGAAGAACACATCAGGCACTGCGCCACCTGCTTCAC
AGTACTTCCCAACAACCTCTTAGAGGTAGGTGTATTCCCGTTTTACAGATAAGGAAACTGAGGC
CCAGAGAGCTGAAGTACTGCACCCAGCATCACCAGCTAGAAAGTGGCAGAGCCAGGATTCAAC
CCTGGCTTGTCTAACCCCAGGTTTTCTGCTCTGTCCAATTCCAGAGCTGTCTGGTGATCACTT
TATGTCTCACAGGGACCCACATCCAAACATGTATCTCTAATGAAATTGTGAAAGCTCCATGTT
TAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQG
RKLDIDFGAEGNRYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNK
LHQQVLWRLVQELCSLKHCFWLERGAGLRVTMHQPVLLCLLALIWLMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

AGGC

no.	address	apartment	city	state	zip
1001	1001	1001	1001	1001	1001
1002	1002	1002	1002	1002	1002
1003	1003	1003	1003	1003	1003
1004	1004	1004	1004	1004	1004
1005	1005	1005	1005	1005	1005
1006	1006	1006	1006	1006	1006
1007	1007	1007	1007	1007	1007
1008	1008	1008	1008	1008	1008
1009	1009	1009	1009	1009	1009
1010	1010	1010	1010	1010	1010
1011	1011	1011	1011	1011	1011
1012	1012	1012	1012	1012	1012
1013	1013	1013	1013	1013	1013
1014	1014	1014	1014	1014	1014
1015	1015	1015	1015	1015	1015
1016	1016	1016	1016	1016	1016
1017	1017	1017	1017	1017	1017
1018	1018	1018	1018	1018	1018
1019	1019	1019	1019	1019	1019
1020	1020	1020	1020	1020	1020
1021	1021	1021	1021	1021	1021
1022	1022	1022	1022	1022	1022
1023	1023	1023	1023	1023	1023
1024	1024	1024	1024	1024	1024
1025	1025	1025	1025	1025	1025
1026	1026	1026	1026	1026	1026
1027	1027	1027	1027	1027	1027
1028	1028	1028	1028	1028	1028
1029	1029	1029	1029	1029	1029
1030	1030	1030	1030	1030	1030
1031	1031	1031	1031	1031	1031
1032	1032	1032	1032	1032	1032
1033	1033	1033	1033	1033	1033
1034	1034	1034	1034	1034	1034
1035	1035	1035	1035	1035	1035
1036	1036	1036	1036	1036	1036
1037	1037	1037	1037	1037	1037
1038	1038	1038	1038	1038	1038
1039	1039	1039	1039	1039	1039
1040	1040	1040	1040	1040	1040
1041	1041	1041	1041	1041	1041
1042	1042	1042	1042	1042	1042
1043	1043	1043	1043	1043	1043
1044	1044	1044	1044	1044	1044
1045	1045	1045	1045	1045	1045
1046	1046	1046	1046	1046	1046
1047	1047	1047	1047	1047	1047
1048	1048	1048	1048	1048	1048
1049	1049	1049	1049	1049	1049
1050	1050	1050	1050	1050	1050
1051	1051	1051	1051	1051	1051
1052	1052	1052	1052	1052	1052
1053	1053	1053	1053	1053	1053
1054	1054	1054	1054	1054	1054
1055	1055	1055	1055	1055	1055
1056	1056	1056	1056	1056	1056
1057	1057	1057	1057	1057	1057
1058	1058	1058	1058	1058	1058
1059	1059	1059	1059	1059	1059
1060	1060	1060	1060	1060	1060
1061</					

148/168

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGT
CEGLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

148/168
MFRSSLLFWPPLCLLSLFLILISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGT
CEGLFYEYIA

FIGURE 149

GTCTCCGCGTCACAGGAACCTTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGAC
TTGACTCCCGCGCGCCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGC
GCCCCAGTCCCGGCCCCCTCTCCCGCCCCACACCCACCCCTCCTGGCTCTTCCTGTTTTTACTCC
TCCTTTTTCATTTCATAACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCC
GAGCGTGGAAGAATGGGGTTTCCTCGGGACCGGCACTTGATTCTGGTGTTAGTGCTCCCGATT
CAAGCTTTCCCCAAACCTGGAGGAAGCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCA
GAAAGACCTTTGAATGAACAGATTGCTGAAGCAGAAGAAGACAAGATTAAAAAACATATCCT
CCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTTTGTGATAACTTGAACCTGCTAAAGGCA
ATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTATAAGAAGCTCCCCACTTGATAAT
AAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAACCTGATCGATGATTATGACTCT
ACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCATCAACTAGACGGG
ACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAGAAAATGAC
AGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGCCAA
GCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAAC
AATTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCA
GAGAAAGTGACTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACA
GTATCTAACACATTAACTTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGAC
AACTTTGAGGAACTCCAATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAA
AAAGAAGCAAAAGAGAAAGAAACACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAG
ATGATGGTGAAATATGGAACAATATCTCCAGAAGAAGGTGTTTCCTACCTTGAAAACCTTGAT
GAAATGATTGCTCTTCAGACCAAAAACAAGCTAGAAAAAATGCTACTGACAATATAAGCAAG
CTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAACAGACAGTACCAAGGAAGAAGCAGCT
AAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAAGATGATAACTCCAACCCAGGA
GGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGCCATCAGAAAAAATATT
GAATGGTTGAAGAAACATGACAAAAAGGGAAATAAAGAAGATTATGACCTTTCAAAGATGAGA
GACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAAGAAGCC
GAGGCCATCAAGCGCATTTATAGCAGCCTGTAAATGGCAAAGATCCAGGAGTCTTTCAAC
TGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATTCTGTGATTAAATTTTTTGACCC
AAGGGTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGC
TTTCTTCCCGTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAA
AAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENK
PGQSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSG
LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTL
EDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMAAIQDGLAKGENDETVSNT
LTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVK
YGTISPEEGVSYLENLDEMIALQTKNKLEKNATDNISKLFPAPEKSHEETDSTKEEAAKMEK
EYGLKDKSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFIN
KQADAYVEKGILDKEEA EAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTC
AAG**ATG**GTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTAT
CTGCATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAG
ATCAGCGTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAG
GGTGAAGCCAGTGCCTGTCATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTG
AACATCATGGAGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTCACCTTCTACCGGCGGGAC
ATGGGGCTCACCTCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCCTGTGCACGGTGCCT
GAAGCCGATCAGCCTGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCCATC
ACAGACTTCTACTTCCAGCAGTGTGACT**TAG**GGCAACGTGCCCCCAGAACTCCCTGGGCAGAG
CCAGCTCGGGTGAGGGGTGAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAG
GACCCCCACGTCTGACTTAGTGGGCACCTGACCACTTTGTCTTCTGGTTCACAGTTTGGATAA
ATTCTGAGATTTGGAGCTCAGTCCACGGTCCTCCCCCACTGGATGGTGCTACTGCTGTGGAAC
CTTGTA AAAA ACCATGTGGGGTAAACTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTG
GGGGAGTGGTGGGAATCATTCCTGCTTAATGGTAACGTGACAAGTGTTACCCTGAGCCCCGCAG
GCCAACCCATCCCCAGTTGAGCCTTATAGGGTCAGTAGCTCTCCACATGAAGTCCTGTCACTC
ACCACTGTGCAGGAGAGGGAGGTGGTGCATAGAGTCAGGGATCTATGGCCCTTGGCCCAGCCCC
ACCCCTTCCCTTTAATCCTGCCACTGTCATATGCTACCTTTCCTATCTCTCCCTCATCATC
TTGTTGTGGGCATGAGGAGGTGGTGATGTCAGAAGAAATGGCTCGAGCTCAGAAGATAAAAGA
TAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAAGATACAATCAAAATCCCAGATGCTGGT
CTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGAAGACCTACTTACAAAGTGGCATATA
TTGCAATTTATTTTAAATTAAGATACCTATTTATATATTTCTTTATAGAAAAAAGTCTGGAA
GAGTTTACTTCAATTGTAGCAATGTCAGGGTGGTGGCAGTATAGGTGATTTTTCTTTTAAATTC
TGTTAATTTATCTGTATTTCTAATTTTTCTACAATGAAGATGAATTCCTTGTATAAAAATAA
GAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGATTGTCTCAGCCTCCAC
TTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGGTTGTAGTAGTGATCA
GGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGTGTGGCTGGAATCT
CTGGGTAAGGAACCTAAAGAACAAAAATCATCTGGTAATTCCTTTCCTAGAAGGATCACAGCCC
CTGGGATTCCAAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGAATTGTGTCC
CCCTCAAATTCACATCCTTCTTGAATCTCAGTCTGTGAGTTTATTTGGAGATAAGGTCTCTG
CAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCAATATGACTG
GTTTCCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTAAAG
ATGAAGGCAGAGATCGGAGTTTTGTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACCA
TCAGAAGCTTGGAAGAGGCAAAGAAGAATTCCTCCCTAGAGGCTTTAGAGGGATAACGGCTCT
GCTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTGGCTGTTT
TAAGCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATC
CCTGTCTCCTCGTGTGTACATTCTGTGTGTGTCCCTCCCACAATGTACCAAAGTTGTCTTTG
TGACCAATAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAGACAC
TGCAGCTTCTACTTGAGCCCTCTCTCTCTGCCACCCACCGCCCCCAATCTATCTTGGCTCACT
CGCTCTGGGGGAAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAA
AATGAAGTCTCCTGCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATC
GATGTTTGTGTGTTTTAAGTTGCTCAGTTTTGGTCTAACTTGTTATGCAGCAATAGATAAATAA
TATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQG
GSQCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPE
ADQPVRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCC
GCCCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTC
TTGGCCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAG
TCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTG
GCTGATAACAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGT
GAGCGCTGCTATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAA
TCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGG
CTAAGCACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAG
GACACAGTGAAAAAGCTTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTG
TTTATGTCTCTGAGAAATGCCTGCATTTGACCCAGAGCAAAGCTGAAAAATGAATAACTAACCC
CCTTTCCTGCTAGAAATAACAATTAGATGCCCCAAAGCGATTTTTTTTAAACCAAAGGAAGA
TGGGAAGCCAAACTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAG
AAACCAATGCCACTTTTGTTTATAAGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTG
ATAACATTTTCATTGTAACCTGGTGTCTATACACAGAAAACAATTTATTTTTTAAATAATTGTC
TTTTTCCATAAAAAAGATTACTTTCCATTCCTTTAGGGGAAAAAACCCTAAATAGCTTCATG
TTTCCATAATCAGTACTTTATATTTATAAATGTATTTATTATTATTATAAGACTGCATTTTAT
TTATATCATTTTATTAATATGGATTTATTTATAGAAACATCATTCGATATTGCTACTTGAGTG
TAAGGCTAATATTGATATTTATGACAATAATTATAGAGCTATAACATGTTTATTTGACCTCAA
TAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEA
SLADNNTDVRLIGEKLFHGVSMSERCYLMKQVLNFTLEEVLFPPQSDRFQPYMQEVVPFLARLS
NRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCT
GTCAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTCAGTGCAGA
GGGCTGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGG
CCAAGCTGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGG
GTGCAG**ATG**AGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAG
GTGGTTGCATTCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGC
CCCAGCAAAGGGCAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCC
CTAGAGCCTGCTAGGCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTC
AACAGCAGGGCCATCTCCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAG
GACCTGTACCACGCCCCGTTGCCTGTGCCCCGCACTGCGTCAGCCTACAGACAGGCTCCCACATG
GACCCCCGGGGCAACTCGGAGCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGC
CATGGCGAGAAGGGCACCCACAAGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTTCCTTA
GCTTGTGTGTGTGTGCGGCCCCGTGTGATGGGC**TAG**CCGGACCTGCTGGAGGCTGGTCCCTTT
TTGGGAAACCTGGAGCCAGGTGTACAACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTT
GGCCCCGTGTGAAGTGCTGTCTGGAGCAGCAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAA
AACCTGCACCTTCTGCACATTTTGAAAAGAGCAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGT
CCTGTCATTTTCTCTCAGGAAAGGTTTTCAAAGTTCTGCCCATTTCTGGAGGCCACCACTCCT
GTCTCTTCCTCTTTTCCCATCCCCTGCTACCCTGGCCCAGCACAGGCACTTTCTAGATATTTT
CCCCTTGCTGGAGAAGAAAGAGCCCCTGGTTTTATTTGTTTGTCTACTCATCACTCAGTGAGC
ATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTTTTGACATGGATGATTCTGAGGAGGA
AGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATATCTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPL
PARPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDP
RGNSELLYHNQTVFYRRPCHGEKGTGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATG**TCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAG
CCGACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTA
ATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTTGCAACAGGGGAC
TATTCAATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAG
GCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAAT
TACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATC
GGCTTCCCTGTAGAGCTGAACACAGTCTATTTTCATTGGGGCCCATAATATTCCTAATGCAAAT
ATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATA
ATGAAATATAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGT
AAGAAGAATGAGGAGACAGTAGAAGTGAACCTCACAACCACTCCCCTGGGAAACAGATACATG
GCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAA
CAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTG
ACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGC
CCACAAACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCT
CTCCTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGG
AGGCACGAAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCATTAAGGTT
CTTGTGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTT
CAAAACCATTGCAGAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGATG
GGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTTCC
AATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTGAGAAC
TCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCCAGATTCAT
CTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATTACAATGCTCTC
AGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTGCAGAACTTCTCCAT
GTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAAGCCTGCCACGATGGCTGCTGCTCCTTG

TAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLI PGDLRLRVEPVTTTSVATGDYS
 ILMNVSWVLRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGF
 PVELNTVYFIGAHNIPNANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKK
 NEETVEVNFTTTPLGNRYMALIQHSTIIGFSQVFEHQKKQTRASVVIPVTGDSEGATVQLTP
 YFPTCGSDCIRHKGTVVLCPTGVFPPLDNNKSKPGGWLPLLLLSLLVATWVLVAGIYLMWRH
 ERIKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFLQNHCRSEVILEKWQKKKIAEMGP
 VQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLAFNLFCSDLRSQIHLH
 KYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVVSAGKRSQACHDGCCSL

Important features of the protein:**Signal peptide:**

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201
 and 283 - 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTG
CTGTTCGATATTGGGGCTTGCCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGA
CATACTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGAC
ATTGGCATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAAACATCGAGAGCCGCTCCACC
TCCCCCTGGAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCC
CAGTGTAGGAACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTT
CCCATCCAGCAAGAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAG
TTGGAGAAGGTGCTGGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAG
TAAGAGGTGCATATCCACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIIN
ENQRVSMERNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQE
TLVVRKHKHGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCA
GGACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGGCTGAAGAGGG
ATTCCAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGG
CAGCACAGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAG**ATG**CCTGTGCCCTGGTTCTT
GCTGTCCCTTGGCACTGGGCGGAAGCCCAGTGGTCCTTTCTCTGGAGAGGCTTGTGGGGCCTCA
GGACGCTACCCACTGCTCTCCGGGCCTCTCCTGCCGCCTCTGGGACAGTGACATACTCTGCCT
GCCTGGGGACATCGTGCCTGCTCCGGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCT
GGTGTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGC
CGTGCATGGGCACTGGGAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGG
GGTGGAGGAGCCTAGGAATGCCTCTCTCCAGGCCCAAGTCGTGCTCTCCTTCCAGGCCTACCC
TACTGCCCCGCTGCGTCTGCTGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTC
TGTGGGCTCTGTGGTATATGACTGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTC
CTATACTCAGCCCAGGTACGAGAAGGAACCTCAACCACACACAGCAGCTGCCTGCCCTGCCCTG
GCTCAACGTGTACAGCAGATGGTGACAACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCA
CTTCGGCCTCTCCCTGTACTGGAATCAGGTCCAGGGCCCCCCTTCCAGGCCTACCC
CCTGACTGGACCGCAGATCATTACCTTGAACCACACAGACCTGGTTCCTGCTCTGTATTCA
GGTGTGGCCTCTGGAACCTGACTCCGTTAGGACGAACATCTGCCCCCTTCCAGGGAGGACCCCCG
CGCACACCAGAACCTCTGGCAAGCCGCCCCGACTGCGACTGCTGACCCTGCAGAGCTGGCTGCT
GGACGCACCGTGTCTGCTGCCCGCAGAAGCGGCACTGTGCTGGCGGGCTCCGGGTGGGGACCC
CTGCCAGCCACTGGTCCCACCGCTTTCCTGGGAGAACGTCACTGTGGACAAGGTTCTCGAGTT
CCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTTCAAGGTGAACAGCTCGGAGAAGCTGCAGCT
GCAGGAGTGTCTGTGGGCTGACTCCCTGGGGCCTCTCAAAGACGATGTGCTACTGTTGGAGAC
ACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCCTTGGAAACCCAGTGGCTGTACTTCACTACC
CAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTACAAGACCTGCAGTCAGG
CCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTGCCCCATGGACAAATA
CATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTGGCCGCTGCGCTTTCCT
CATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAACAGGACGTCCG
CTCGGGGGCGGCCGCCAGGGGGCGCGGGCTCTGCTCCTCTACTCAGCCGATGACTCGGGTTT
CGAGCGCCTGGTGGGCGCCCTGGCGTCCGCCCTGTGCCAGCTGCCGCTGCGCGTGGCCGTAGA
CCTGTGGAGCCGTCTGAACTGAGCGCGCAGGGGGCCCGTGGCTTGGTTTACGCGCAGCGGCG
CCAGACCCTGCAGGAGGGCGGCGTGGTGGTCTTGTCTTCTCTCCCGGTGCGGTGGCGCTGTG
CAGCGAGTGGCTACAGGATGGGGTGTCCGGGGCCCGGGGCGCACGGCCCGCACGACGCTTCCG
CGCCTCGCTCAGCTGCGTGTGCTGCCCGACTTCTTGCAGGGCCGGGCGCCCGGCAGCTACGTGGG
GGCCTGCTTTCGACAGGCTGCTCCACCCGGACGCGGTACCCGCCCTTTTCCGCACCGTGCCCGT
CTTACACTGCCCTCCCAACTGCCAGACTTCTTGGGGGCCCTGCAGCAGCCTCGCGCCCCGCG
TTCCGGGGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCTGGGGCCCTTCCAGCCAGCCCTGGATAG
CTACTTCCATCCCCCGGGGACTCCCGCGCCGGGACGCGGGGTGGGACCAGGGGCGGGACCTGG
GGCGGGGGACGGGACT**TAA**ATAAAGGCAGACGCTGTTTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRVLVGPQDATHCSPGLSCRLWDSILCLPGDIVPAPGPVLAP
THLQTELVLRQCKETDCDLCLRVAVHLAVHGHWEPEDEEEKFGGAADSGVEEPRNASLQAQVV
LSFQAYPTARCVLLEVQVPAALVQFGQSVGSSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQ
QLPALPWLNVSADGDNVHLVLNVSEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDL
VPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLTLQSWLLDAPCSLPAEAALCW
RAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQVNSSEKLQQLQECLWADSLGPLKD
DVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQSGQCLQLWDDDLGALW
ACPMDKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSAAAARGRAALLLY
SADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFWFAQRRQTLQEGGVVLLFS
PGAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLHPDAVPA
LERTVPVFTLPSQLPDFLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPGTPAPGRGV
GPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394**Glycosaminoglycan attachment site.**

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCG**ATG**AGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGG
CTGCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACT
TTGAAAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCG
AGTATAAGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGA
AGTCCTGCAACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTTCAGCTCTCTGCAGCACACT
ACCCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTTCAGATGATTGTTTCAT
CCTACCCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCAT
GACCTGTTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAG
CAGAGAGAATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTTCCTTGGCACCATCATGATT
TGCGTTCCACCTGGGCCAAGGAGAGTGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGAC
CGGACATGGACCTACTCCTTCTCCGGAGCCTTCCTGTTCTCCATGGGCTTCCTCGTCGCAGTA
CTCTGCTACCTGAGCTACAGATATGTACCAAGCCGCCTGCACCTCCCAACTCCCTGAACGTC
CAGCGAGTCCTGACTTTCCAGCCGCTGCGCTTCATCCAGGAGCACGTCCTGATCCCTGTCTTT
GACCTCAGCGGCCCCAGCAGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGA
CCCAGGGAGCCCCGAGGAGCTCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAG
CCAGACATCTCCATCCTCCAGCCCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCC
TATGCCCCAAACGCTGCCCCCTGAGGTGCGGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAA
GCTCAATTCCCATTTCTACGCCCCACAGGCCATCTCTAAGGTCCAGCCTTCCTCCTATGCCCCCT
CAAGCCACTCCGGACAGCTGGCCTCCCTCCTATGGGGTATGCATGGAAGGTTCTGGCAAAGAC
TCCCCCACTGGGACACTTTCTAGTCTTAAACACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAG
CCACCAGCTGGAAGCTGCATGTTAGGTGGCCTTTCTCTGCAGGAGGTGACCTCCTTGGCTATG
GAGGAATCCCAAGAAGCAAAATCATTGCACCAGCCCCCTGGGGATTTGCACAGACAGAACATCT
GACCCAAATGTGCTACACAGTGGGGAGGAAGGGACACCACAGTACCTAAAGGGCCAGCTCCCC
CTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTCCCTCCCTTTGCAACCTCCTTCCGGT
CCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCCTGCTGGAGTCCCTTGTGTGTCCC
AAGGATGAAGCCAAGAGCCCAGCCCCCTGAGACCTCAGACCTGGAGCAGCCCACAGAACTGGAT
TCTCTTTTCAGAGGCCTGGCCCTGACTGTGCAGTGGGAGTCC**TG**AGGGGAATGGGAAAGGCTT
GGTGCTTCCTCCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCAATCCCATGCCTGCCCCAT
GCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGCAGAGGGAGTGGCATG
CAGGGCCCCCTGCCATGGGTGCGCTCCTACCGGAACAAAGCAGCATGATAAGGACTGCAGCGG
GGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTGCAAGGCAGAAA
TGACAGTGCAAGGAGGAAATGCAGGGAAACTCCCGAGGTCCAGAGCCCCACCTCCTAACACCA
TGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTCACAATCT
AGCTCGACAGAGCATGAGGCCCCCTGCCTCTTCTGTCAATTGTTCAAAGGTGGGAAGAGAGCCTG
GAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGCACT
TCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCAT
TCCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTTCAGCTTCATTCTCTGATAGAACAAAG
CGAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGAC
CCTATCCTGAGAATGGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAA
CACACTGTACTGATGTCACAACCTTTGCAAGCTCTGCCTTGGGTTTCAGCCCATCTGGGCTCAAA
TTCCAGCCTCACCCTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCCAAGAACCTC
GGTTTCTCTCATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAA
ATGAAGTCATGTCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGT
AGCTATTTAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGER
DWVAKKGCQRITRKSCNLTVETGNLTELYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDV
TCISKVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRITYQMHLGGKQREYEFF
GLTPDTEFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYR
YVTKPPAPPNSLVQRVLTFQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGA
PQRHSLSEITYLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYA
PQAISKVQPSSYAPQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCM
LGGLSLQEVTSLAMEESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLKGQLPLLSSVQI
EGHPMSLPLQPPSGPCSPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLA
LTVQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCCGCGGTGGCCACAACAT
GGCTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGG
CCAGTCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTG
CAGCATGTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGT
GAATTTTAAAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGA
TTGGGCTGGAAGTGTGAACACAGTTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCA
TAAATACACGGAAGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGG
AGGAAGAGATGATTTTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGA
CTCTGTACCTGAAGAGTCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGA
GGAGTCTCGGGGGCGTGAACCTTGACCCTGTGCCTGAGCCCGAGGCATTCAGAGCTGATTCAGA
GGATGGAGAAGGTGCTTTCTCAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGA
GAGCCACCCTCACACCAGCGGTCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACAC
TTTTGAAGAAATTCTGCACGATAAATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAG
TTCTCCTGCCTCGGTGGAGCGGGAGAAGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAG
TCAGAGAGGAAGTGGACAGTGC GTTATT CATTACAGCAAAGGATTTTCGTTGGCATCAAATCT
AAGTTTGTTTTACAAAGATTGTTTTTTAGTACTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCC
AAACAAAATATATTATTTTCCCTTCTAAGTAAAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWL FVLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPDCRF
VNFKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEELHIPADETDFVCFE
GGRDDFNSYNVEELLGSLELED SVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRADS
EDGEGAFSESTEGLOGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGN
SSPASVEREKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTGAGAGGCCGGGGAAGAGAAGCAAAGC
GCAACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAA
CTTCAGTCCCCCAAACGCGCACCCCTCGAAGTCTTGAAGTCCAGCCCCGCACATCCACGCGCGG
CACAGGCGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTTCGGGCAGCTGG
GCTCGGGCGGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGG
CTGCGCCCTGGGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGA
TGAGCCGCGTGGTCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCC
GCCGCGTGGTCAAGCGGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGG
CCTACTTCCATGAAGTGTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTG
AGGGAGGAGTCCTCCTCAGCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGC
AAAACCTGACAAAACCCGGGACAGGGATTCTGATGGTGATTCTGGATAGGGCTTTGGAGGA
ATGGAGATGGGCAAACATCTGGTGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATT
CCAGTACCGAACTGGTACACAGATGAACCTTCTGCGGAAGTGAAAAGTGTGTTGTGATGT
ATCACCAACCAACTGCCAATCCTGGCCTTGGGGGTCCCTACCTTTACCAGTGAATGATGACA
GGTGTAAACATGAAGCACAATTATTTGCAAGTATGAACCAGAGATTAATCCAACAGCCCCTG
TAGAAAAGCCTTATCTTACAAATCAACCAGGAGACACCCATCAGAATGTGGTTGTTACTGAAG
CAGGTATAATTCCCAATCTAATTTATGTTGTTATACCAACAATACCCCTGCTCTTACTGATAC
TGGTTGCTTTTGGAACTGTTGTTTCCAGATGCTGCATAAAAGTAAAGGAAGAACAACAACTA
GTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTACCAGAAAAGAAAGTGGCATGGAAGTAT
AATAACTCATTGACTTGGTTCCAGAATTTTGTAAATTCTGGATCTGTATAAGGAATGGCATCAG
AACAAATAGCTTGAATGGCTTGAAATCACAAAGGATCTGCAAGATGAACTGTAAGCTCCCCCT
TGAGGCAAATATTAAAGTAATTTTTATATGTCTATTATTTTCAATTTAAAGAATATGCTGTGCTA
ATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAACTTCAAACCTTCAAGCAAA
TGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTATGTGTGTTAGAAGCAAT
TCCTTTTATTTCTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAATGTATATTGTATTGA
AATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGATAAAAATGAACTGTTCTA
ATATTTATTTTTATGGCATCTCATTTTTTCAATACATGCTCTTTTGATTAAAGAACTTATTAC
TGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGTGTTTCTCGAA
ATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGAAATAAGA
AGCTATTTTATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGATTGTCT
AATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTGTAGCTTAAATTTAAACAGATTTTG
TAATAATGTAACCTTGTAAATAGGTGCATAAACTAATGCAGTCAATTTGAACAAAAGAAGT
GACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTG
AGGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTT
GGGATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGC
CTCTGACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAG
TGTGGAGACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACTACGAAATCGTGTGAAA
ATGGGTTGGAACCCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATG
GTGGCTCCTTTCTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTTCTTCTCAAGAGA
AAGTTGTAACCTCTCTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAATAAAGAGTTCTTG
TTTCTGGGGGAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVVSQKVCFAFDKHPCKMAYFHELSSRVSFQEARLACES
EGGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDWIGLWRNGDGQTSGACPDLYQWSDGSN
SQYRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAP
VEKPYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSCKGRTKT
SPNQSTLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217